*Final-text*JEE-0331

Running heads: recto Student Response to Instructional Practices

Verso: DeMonbrun, Finelli, Prince, Borrego, Shekhar, Henderson, & Waters[e]

Files:
figures
captions
text
tables

&&&&&&&&&&&&&&&&&&

Creating an Instrument to Measure Student Response to Instructional Practices

Matt DeMonbrun[a], Cynthia J. Finelli[a], Michael Prince[b], Maura Borrego[c],
Prateek Shekhar[a], Charles Henderson[d], and Cindy Waters[e]

[a]University of Michigan, [b]Bucknell University, [c]University of Texas, Austin,
[d]Western Michigan University, [e]North Carolina A&T State University

(1)Abstract

Background Calls for the reform of education in science, technology, engineering, and mathematics (STEM) have inspired many instructional innovations, several of which are research based. Yet the adoption of such types of instruction has been slow. Research has suggested that students' response may have a significant effect on instructors' willingness to adopt different types of instruction.

Purpose We created the Student Response to Instructional Practices (StRIP) instrument to measure the effects of several variables on student response to instructional practices. We discuss the step-by-step process for creating this instrument, from the initial development through multiple stages of validity and reliability testing.

Design/Method The development process had six steps: item generation and construct development, validity testing, implementation, exploratory factor analysis, confirmatory factor analysis, and instrument modification and replication. We discuss the pilot testing of the initial instrument on 362 students, and developing the constructs and validation using exploratory and confirmatory factor analyses.

Results This process produced 47 items measuring three parts of our framework. Types of instruction separated into four factors (interactive, constructive, active, and passive); strategies for using in-class activities into two factors (explanation and facilitation); and student responses to instruction into five factors (value, positivity, participation, distraction, and evaluation).

Conclusions This study describes the design process and final results for our instrument, a useful tool for understanding the relationship between the type of instruction used and students' response.

Keywords active learning; instructional methods; factor analysis; student resistance

(1)Introduction

There have been various calls for the reform of education in science, technology, engineering, and mathematics (STEM), including increasing the number and diversity of students receiving these degrees (AAAS, 2010; NAS, NAE, & IOM, 2007). These calls for reform have drawn forth many innovations in the types of instruction used in the classroom, several of which are research based (Jamieson & Lohmann, 2012; Kuh, 2008; Seymour & Hewitt, 1997). Many of these research-based types of instruction fall under the broad definition of active learning, or requiring students to participate in class activities beyond watching an instructor lecture (Felder & Brent, 2009); prior research has shown active learning can be especially effective for educating a diverse student body (Prince, 2004; Seymour & Hewitt, 1997) and for increasing the retention rate of students in STEM programs (Angelo & Cross, 1993; Prince & Felder, 2006).

Despite this literature base, translation of research about innovative types of instruction to instructional practice has been slow (Friedrich, Sellers, & Burstyn, 2007; Handelsman et al., 2004; Hora, Ferrare, & Oleson, 2012; PCAST, 2012; Singer, Nielsen, & Schweingruber, 2012). Several researchers have identified a number of instructor-reported barriers that help to explain these slow adoption rates. Among the least researched but most often mentioned barriers is the concern that students will resist, or respond in negative ways (Borrego, Froyd, & Hall, 2010; Dancy & Henderson, 2012; Finelli, Daly, & Richardson, 2014; Froyd, Borrego, Cutler, Prince, & Henderson, 2013; Henderson & Dancy, 2007; Seidel & Tanner, 2013). In actuality, student response to new and different types of instruction can be positive if students are engaged in these activities, view them in a positive light, and see the value in their use (Gauci, Dantas, Williams, & Kemm, 2009; Livingstone & Lynch, 2000). However, worries about such negative responses can discourage instructors from adopting new and different types of instruction.

Research that characterizes the types of student response (both positive and negative) to various types of instruction and identifies strategies for introducing these types of instruction could help eliminate a key barrier to faculty adoption new instructional practices. And although literature offers a variety of tips for instructors wishing to promote positive response and minimize negative reactions to different types of instruction (e.g. Armstrong, 1998; Arum & Roksa, 2011; Felder, 2011; Johnson, Johnson, & Smith, 1991; Lake, 2001; Michael, 2007; Moffett & Hill, 1997; Prince, Borrego, Henderson, Cutler, & Froyd, 2013), these suggestions tend to be drawn from personal experience and have yet to be empirically tested. These limitations show the need for additional research in this area. Such research requires an instrument to assess and measure students’ responses to different types of instruction and the strategies used (or not used) with each instructional type. Here, we report on development of the Student Response to Instructional Practices (StRIP) instrument to achieve this goal.

Rather than focusing on the effects of instructional types, this article describes the development process of the StRIP instrument, which followed accepted approaches for instrument development (e.g., Carberry, Lee, & Ohland, 2010; Li, McCoach, Swaminathan, & Tang, 2008; Ro, Merson, Lattuca, & Terenzini, 2015). The resulting StRIP instrument can be used by researchers and practitioners seeking a tool to study student response to all types of instruction in the classroom, and the framework we have developed attempts to explain the relationship between types of instruction, strategies for using these types of instruction during class, and how students respond.

(1) Methods

We adapted the development process for the Student Response to Instructional Practices (StRIP) instrument from Hinkin (1998), as shown in Figure 1. The process is iterative, which involves a six-step approach:

1. Generating items and developing constructs for the instrument; the process borrows from prior literature on instructional types, student response, and strategies for using in-class activities,

2. Testing for validity by observing the engineering classroom, as well as by subjecting the instrument to expert review and cognitive interviewing,

3. Implementing the instrument,

4. Conducting an exploratory factor analysis, an important step since there was no previously instrument on student response to instructional practices,

5. Conducting a confirmatory factor analysis to verify the constructs established in the exploratory factor analysis, and

6. Modifying the instrument and replicating findings through full instrument administration.

[Figure 1 here]

(2)Step 1: Item Generation and Construct Development

In creating a new instrument, researchers must first generate the items needed to measure the desired construct(s), a process that can be accomplished through deductive or inductive scale development. Given the limited amount of empirical research and absence of a developed framework on students' responses to types of instruction, we chose an inductive approach to item generation (Ironson, Smith, Brannick, Gibson, & Paul, 1989). Figure 2 represents the framework we developed in order to better understand students' responses to types of instruction; the framework comprises several groups of variables that potentially contribute to student response. Instructors likely influence student response by their choice of instructional strategy (e.g., lecturing or active learning) and how they introduce and manage that strategy in the classroom. We hypothesized student response depends in part on student characteristics, preferences, expectations, and prior experiences. The framework features characteristics of the course itself and clarifies that a student's reason for taking the course potentially influences his or her response to types of instruction.

The three sections of the instrument correspond to the three parts of our framework:

Types of instruction

Strategies for using in-class activities

Student responses to instruction

Types of instruction Because students' responses vary according to the types of instruction experienced by the student, we developed items to capture these instructional types, ranging from traditional lecture to simple and more complex forms of active learning. While trying to characterize these types of instruction by the nature of what occurs during the instruction, such as individual work, group work, and pair and share, we also wanted to frame them around the types of cognitive processes used by students during the activities to understand whether or not certain types of instruction shape students' responses.

[Figure 2 here]

We modified Chi and Wylie's (2014) interactive-constructive-active-passive (ICAP) model, which classifies instructional activities as interactive, constructive, active, or passive learning processes. Although our modified version uses the same format as the original ICAP, we redefined some of the original terminology to be more consistent with other research on active learning (Felder & Brent, 2009; Prince, 2004). We made three modifications.

First, we sought to differentiate between active and passive types of instruction. Both types of instruction involve the individual students' actions (or lack thereof) during the instructional practice. We defined *passive instruction* as occurring when students are expected to passively receive information from the instructor. Examples include listening to lectures or watching the instructor solve problems on the board. Since our focus for passive instruction is on information received directly from the instructor, we did not include textbooks and other resources when asking students about information sources. We defined *active instruction* as occurring when students are engaged with the course content in any individual activity. Examples include asking the instructor questions or answering questions posed by the instructor during class.

Because there is clear evidence that team and group activities can generate high levels of negative student response (Bacon, Stewart, & Silver, 1999; Donohue & Richards, 2009; Lake, 2001; Oakley, Hanna, Kuzmyn, & Felder, 2007; Powell & Kalina, 2009), we made a distinction between individual activities and those with two or more students. For the latter we used the term *interactive instruction*, which is similar to Chi and Wylie's (2014) use of the term. Our conceptualization differs, however, in that we included any interaction students might have with their peers during the semester (including studying or completing homework in groups), while Chi and Wylie (2014) stipulate that the activity must involve students creating knowledge together; for example, students must have a dialogue with other students. Examples of interactive types of instruction include doing hands-on group activities during class and being graded based on the performance of a group.

Finally, some complex types of active learning include elements such as self-directed learning and ill-structured problems that have been hypothesized to generate significant student resistance (Hung, Bailey, & Jonassen, 2003; Van Barneveld & Strobel, 2011; Yadav, Subedi, Lunderberg, & Bunting, 2011). These types of activities are defined by learning on one's own (self-discovery), rather than learning from being told what to do (direct instruction; Chi, 2009). Thus, we retained Chi and Wylie's (2014) definition for *constructive instruction* for these instructional types since they place high expectations on students and represent significant departures from many traditional classes.

Altogether, we created 21 items for students to report the frequency of these types of instruction (Table 1). We also asked students whether they wanted more or fewer of these activities in their ideal course to gauge their desired frequency. We expected students' responses to a particular type of instructional practice to be based not only on the actual level of use, but also on the difference between the actual and desired levels of use.

Strategies for using in-class activities While little empirical research has investigated the effectiveness of strategies for using in-class activities, several authors give advice about how to introduce different types of instruction and minimize negative reactions (Armstrong, 1998; Bentley, Kennedy, & Semsar, 2011; Moffett & Hill, 1997; Van Barneveld & Strobel, 2011). We included these strategies in the StRIP instrument to allow more thoughtful analysis of their relative effectiveness.

[Table 1 here]

Three themes emerge from the literature on reducing student resistance. First, beginning the course activity with an explanation of its purpose and process and an acknowledgment of its challenges can better prepare students for what is expected of them and why the activity is important (Bacon et al., 1999; Yadav, Subedi, Lunderberg, & Bunting, 2011), especially if their participation might affect their grade (Donohue & Richards, 2009). Indeed, Gaffney, Gaffney, and Beichner's (2010) Pedagogical Expectancy Violation Assessment acknowledges that students' expectations of active learning can fluctuate throughout the semester, and that this fluctuation can affect students' responses to the activities. Second, soliciting student feedback and providing the support needed to successfully complete the activity assists students in achieving their goals (Bentley et al., 2011; Yadav et al., 2011). Finally, designing appropriately challenging activities ensures that all students can successfully attempt and complete the activity (Donohue & Richards, 2009; Van Barneveld & Strobel, 2011).

We used both the published strategies suggested for using in-class activities and strategies we observed in our prior research (Shekhar, DeMonbrun, et al., 2015) as we developed the strategies for using in-class activities items on the StRIP instrument. Altogether, we created eight items for students to report how frequently the instructor engaged in the recommended strategies (Table 2).

[Table 2 here]

Student responses to instruction To characterize students' responses to types of instruction, we drew upon ideas found in the literature, including the school classroom engagement concept of Fredricks, Blumenfeld, and Paris (2004), Chasteen's (2014) construct of productive engagement, and Weimer's (2002) framework on student resistance. The idea of engagement is often characterized as the responses students have to their experiences at specific moments in time (Lawson & Lawson, 2013). Such responses can range from moments of total engagement or flow to more passive moments of boredom or lack of interest (Pekrun & Linnebrink-Garcia, 2012). Hence, we designed our instrument to examine how types of instruction facilitate students' engagement in the classroom, but we also wished to address faculty concerns regarding student resistance to these types of instruction, rather than simply measure boredom or lack of engagement.

Previous research has conceptualized three forms of classroom engagement: cognitive engagement (psychological investment in classroom activities), affective-emotional engagement (social and emotional connections to the classroom), and behavioral engagement (students' behavior in the classroom; Appleton, Christenson, & Furlong, 2008; Fredericks et al., 2004; Furlong & Christenson, 2008; Skinner & Pitzer, 2012). To these three forms of engagement we added a fourth concept of evaluation, because of the value instructors place on end-of-semester student ratings. We constructed four subscales:

value – the degree to which students see the activity as worthwhile (cognitive);

positivity – how positive or negative students feel about the activity (affective-emotional);

participation – the extent to which students do or do not participate or demonstrate resistance (behavioral); and

evaluation – the way students rate the instructor or course at the end of the term.

Value Chasteen (2014) defines value as a measure of some elements of cognitive engagement that are affected by students' thoughts, beliefs, and expectations. In their review of school engagement, Fredricks et al. (2004) indicated that cognitive engagement stresses students' investment in their learning and incorporated literature on learning and instruction, self-regulation, and investment in learning. There are several conceptualizations of cognitive engagement, which include a desire to go beyond the typical requirements of a course (Connell & Wellborn, 1991; Newmann, Wehlage, & Lamborn, 1992; Wehlage, Rutter, Smith, Lesko, & Fernandez, 1989) and a self-regulated motivation to learn and do well in a course (Brophy, 1987; Pintrich & De Groot, 1990; Zimmerman, 1990). In our instrument, value is related to students' investment in their learning. At the high end of the value scale, students understand and accept the rationale for the activity, and they feel the time used for the activity is beneficial. At the other end of the scale, students tend to disagree with the rationale for the activity and feel that time could be better spent doing other things.

Positivity Affective-emotional engagement refers to the affective reactions of students in the classroom, including anxieties, feelings of belongingness, happiness, sadness, interest, and boredom (Connell & Wellborn, 1991; Skinner & Belmont, 1993). Although a traditional scale of academic emotions (Pekrun, Goetz, Titz, & Perry, 2002) measures how students' goals affect their own emotions in the classroom setting (Lee & Smith, 1995; Pekrun, Elliot, & Maier, 2009;

Stipek, 2002), the context of our StRIP instrument is different in that it measures students' reactions to the instructor and the course. Thus, we decided to label this factor as *positivity* to avoid any confusion with the academic emotions scale. At the high end of this scale, students feel positively about the task, instructor, and classroom environment. Students with low positivity respond in a negative way.

Participation Because the research on behavioral engagement is considerably broad (Lawson & Lawson, 2013) and often captures student behavior outside of the classroom (Finn, Folger, & Cox, 1991), we opted to constrain behaviors in our instrument to those exhibited only in the college classroom. Chasteen's (2014) work provided guidance for the positive components of behavioral engagement included in the instrument; we applied Weimer's (2002) framework on student resistance to further distinguish the negative components. Weimer identified three types of resistance, or negative behavioral engagement:

> **Open resistance** On some occasions, students openly object to the approach. They may demonstrate open resistance by complaining, arguing, or objecting, and they generally do so in ways that are not constructive.
>
> **Passive, nonverbal resistance** Students exhibit an overall lack of enthusiasm as a way to assert their objection to the approach. Students may demonstrate passive, nonverbal resistance by not doing assignments but offering excuses, faking attention, or appearing to take notes while working on material from another class.
>
> **Partial compliance** Students may demonstrate partial compliance by completing a task poorly, half-heartedly, or quickly, by putting forth minimal effort, or by being preoccupied with procedural details.

We labeled this factor *participation*. The items on our StRIP instrument in the participation subscale represent both these positive and negative components of participation.

Evaluation Another significant element of students' responses is evaluation, or how students rate both the overall course and quality of instruction on course evaluation forms. Since student evaluations play a significant role in many instructors' retention, tenure, and promotion reviews, low student ratings are clearly an important response that is likely to influence whether instructors adopt and continue to use various types of instruction in their classes. To capture this element of students' responses, we added items to our StRIP instrument about the quality of the course and its instruction. These items were based on similar items from the Individual Development and Educational Assessment (IDEA)student survey form (Cashin, 1988, 1990).

Altogether, we created 15 items for students to report how often they responded in various ways to the types of instruction that were used in their course. These items are listed in Table 3.

[Table 3 here]

(2)Step 2: Validity Testing

In our second step, testing for validity, we wanted to ensure that the proposed uses for the instrument were appropriate given the context and purposes of our study (AERA, APA, & NCME, 2014). Specifically, we developed our StRIP instrument to measure students' responses to types of instruction encountered in the undergraduate engineering classroom. Therefore, the process of establishing evidence for the validity of our measures was achieved in a number of ways: using multiple, mixed-methods approaches for development and validation (Haynes, Richard, & Kubany, 1995); subjecting the instrument to expert review (Nunnally & Bernstein, 1994); conducting cognitive interviewing with potential respondents of the instrument (Nunnally & Bernstein, 1994); and reporting results to expert reviewers (Hinkin, 1998). We especially used classroom observations, expert review, and cognitive interviewing during this validation process. These are all standard practices for establishing validity as according to the *Standards for Educational and Psychological Measurement* (AERA, APA, & NCME, 2014). As indicated by the recursive nature of steps 1 and 2 in Figure 1, this process often led to generating new items and revising factors based on feedback from these various sources.

Classroom observations In addition to our extensive literature review and item development process, we recognized the need to collect more concrete data about students' responses to types of instruction. We conducted classroom observations to inform the instrument development process. During our survey development process, we conducted observations in four large introductory engineering courses, ranging in size from 70 to 150 students, at two large public research universities (Shekhar, DeMonbrun, et al., 2015).

These observations served three purposes. First, by collecting first-hand observations of various types of students' responses to instruction, we further confirmed our framework (Figure 1). Second, we observed strategies for using in-class activities that were not mentioned in the literature and which we subsequently added to our instrument. Specifically, we included two items from Table 2 ("Used activities that were the right difficulty level (not too easy, not too difficult)" and "Walked around the room to assist me or my group with the activity, if needed") to address strategies observed in the classroom. Finally, we pilot tested the StRIP instrument in some of the same classes we observed; this testing allowed us to study the extent to which students' responses about types of instruction were related to our independent observations. Using these observations as a form of triangulation (Greene, Caracelli, & Graham, 1989), we gained confidence in the instrument's ability to measure the underlying factors in our study.

Expert review Following our initial review of the literature, we created a preliminary draft of the StRIP instrument and invited our three-member advisory board to offer their expert critique. The board included faculty who were experienced in instrument design and psychometrics, types of instruction, and students' responses to different types of instruction. Their feedback aided in refining our instrument. They provided guidance on timing and logistics for implementing the instrument, suggested that we find a framework for our instructional types and include items related to positivity and enjoyment, and recommended we clarify the response scale for the items by incorporating Fraser's (1998) classroom environments frequency scale rather than using a typical Likert scale response.

Cognitive interviewing Following the approach used by Ouimet, Bunnage, Carini, Kuh, and Kennedy (2004), we conducted cognitive interviews (Willis, 2004) with 12 undergraduate engineering students at three institutions to confirm that the instrument was well designed for the target audience. We asked students to review each individual item; describe what they thought the item was asking, how they would respond, and how they would arrive at their response; and talk about other issues such as clarity of items and response scales and ease of completion. These cognitive interviews provided assurance that the students' interpretations of the instrument and its individual items were aligned with the intended constructs. Student feedback allowed us to better organize the instrument and reformat some question prompts. Specifically, these students suggested that we move the student responses to instruction section to the front of the instrument, because it allowed them to think broadly about their experiences in class before outlining specific practices in the types of instruction section.

(2)Step 3: Implementation

Next we pilot tested the draft instrument in two phases. During the first phase, we studied 191 students in four courses from three institutions; during the second phase, we studied an additional 171 students in four courses from three institutions. Across both phases, we administered the instrument to a total of 362 students in eight courses at four institutions. Additional information on the courses in our sample in given in Table 4.

[Table 4 here]

We selected courses for our pilot testing through a mix of convenience and purposive sampling (Teddlie & Yu, 2007). A member of our research team at each of four institutions chose one or two instructors teaching gateway engineering courses on the basis of their prior knowledge of their instructional methods. All students in those classes were asked to complete the StRIP instrument. Although students were offered an opportunity to opt out of taking the instrument, we are not aware of any students who did so. Therefore, no sample weights were used, because our selection was representative of each course. Only 11 responses had missing or incomplete data on any of the items. Because this number was less than 3% of the total sample and the missing data pattern appeared to be random (Rubin, 1976), these surveys were removed from the analyses. We used data from the first phase of pilot testing for an exploratory factor analysis and the second phase for confirming the factors identified in the first phase. All analyses were performed using Stata 13.1 SE software.

(2)Step 4: Exploratory Factor Analysis

We conducted an exploratory factor analysis (EFA) on the StRIP instrument to identify emergent factors from our first phase of pilot testing and to determine items that might be particularly problematic given low or multiple factor loadings. The EFA included 191 responses to 44 items, giving us a 4:1 ratio of respondents to items, remaining above recommendations for a 3:1 participant-to-item ratio (Reise, Waller, & Comrey, 2000; Thompson, 2004).

Because we were studying three categories of variables, we conducted three separate EFAs. Using a common-factors method and promax oblique rotation (recommended for intercorrelated measures by Worthington and Whittaker, 2006), we identified four factors for types of instruction, two factors for strategies for using in-class activities, and four factors for student responses to instruction (as described subsequently and shown in Table 3, we later split this construct into five factors). The factors and their loadings are also listed in Tables 1, 2, and 3. All factors had eigenvalues above 1.0 (Kaiser, 1958), and each EFA model was tested using standard tests of significance (Bartlett's test of sphericity) and sampling adequacy (Kaiser-Meyer-Olkin). All models were statistically significant ($p < 0.001$), indicating that the variables were intercorrelated, and their sampling adequacies were above the 0.60 required for good factor analyses (Tabachnick & Fidell, 2001). All items had a loading at or above the threshold of 0.32 (Comrey & Lee, 1992), and each construct had a construct reliability above the recommended benchmark of 0.60 (Bagozzi & Yi, 1988). For the evaluation construct, we used the Spearman-Brown coefficient to measure construct reliability, as recommended in previous research (Eisinga, Grotenhuis, & Pelzer, 2013).

Based on the response loadings in each EFA, we developed a name for each factor to assist in describing the phenomenon captured by the groupings. For types of instruction (Table 1), we conducted an EFA on both students' ideal types of instruction as well as what they actually experienced in the course. While the factors related to ideal instruction closely aligned to our adaptations of the ICAP framework (Chi & Wylie, 2014), those related to actual experience did not. We hypothesize this occurred because, while students tend to think about ideal types of instruction in term of the interactive, constructive, active, and passive categories, the capabilities of an instructor to balance each of these types in actual instruction might be limited. Therefore, we only present the analyses for the ideal types of instruction.

For strategies for using in-class activities (Table 2), we identified factors including explanation strategies (where the instructor was the main character in the strategy and took the role of explaining the activity) and facilitation strategies (where the instructor facilitated opportunities for students to participate in the strategy). For student responses to instruction (Table 3), although we initially designed the instrument with four subscales, the EFA resulted in two factors that emerged from the participation factor – student distraction and student participation. Distraction contains items where students distract themselves or peers during the learning process, whereas participation indicates the extent to which students participated in the activity. All five resulting factors and their loadings are presented in Table 3.

(2)Step 5: Confirmatory Factor Analysis

Given the success of the initial pilot testing, we used data from the second pilot phase of the StRIP instrument to conduct a confirmatory factor analysis (CFA); see Tables 5, 6, and 7 to verify the reliability of the factors. The CFA included 171 responses to 44 items giving us a nearly 4:1 ratio of respondents to items, which falls within recommended minimum sample size (Kline, 2005). The purpose of the CFA was to test the model identified in the EFA for structural fit to the developed constructs. Recently, researchers have turned to structural equation modeling (SEM) rather than standard factor analysis techniques to conduct a CFA (Martens, 2005; Martens & Hasse, 2006). Usually, SEM consists of two steps in the model-building process: testing for the factorial validity of a theoretical construct (first-order CFA model) and a path analysis to describe the relationship between theoretical constructs. Given our desire to replicate the latent factors of the instrument, as opposed to determining their relationship(s) with other factors, we chose to only conduct a first-order CFA model (Byrne, 2013).

The test statistics indicated good overall model fit. The chi-square statistic for the model was 2.98, falling below the recommended threshold (Kline, 1998). The root-mean-square error of approximation was 0.06, with the lower bound of our 90% confidence interval at 0.00 and an upper bound at 0.14, suggesting a reasonable fit to the model. The comparative fit index statistic was 0.98, indicating good model fit (Hu & Bentler, 1999). Finally, the standardized root-mean-square residual was 0.03, considered to be favorable for the model (Hu & Bentler, 1999).

In addition to the factor loadings, we also display the standard error, item reliability, average variance extracted, and construct reliability of each of the factors in Tables 5, 6, and 7. Item reliabilities ranged from 0.51 to 0.89, which exceed the acceptable value of 0.50 (Hair, Anderson, Tatham, & Black, 1992). The average variance extracted for all constructs was well above the threshold value of 0.5 (Fornell & Larcker, 1981). Finally, the reliabilities for each construct were above the benchmark of 0.60 (Bagozzi & Yi, 1988). As noted in the Exploratory Factor Analysis section, the construct reliability for our two-item evaluation construct was conducted using the Spearman-Brown statistic, as is recommended with the use of two-item scales (Eisinga et al., 2013).

[Table 5 here]

[Table 6 here]

[Table 7 here]

(2)Step 6: Instrument Modification and Replication

After conducting the exploratory and confirmatory factor analyses, we engaged in instrument modification and replication to further strengthen the instrument. In the EFA and CFA, we found that two of the student responses to instruction items loaded strongly on two different factors. These items included “I pretended but did not actually participate” and “I rushed through the activity, giving minimal effort.” We determined this instance of double-loading to be the result of items being worded as compound statements. The statements “I did not actually participate” and “I gave minimal effort” appeared to relate to the participation factor (the standardized factor loadings from our CFA were 0.71 and 0.64, respectively), while the statements “I pretended to participate” and “I rushed through the activity” appeared to relate to distraction (the standardized factor loadings from our CFA were 0.70 and 0.63, respectively). Therefore, we split these items to create four items to address both of the factors:

> I did not actually participate in the activities (participation)
>
> I gave the activities minimal effort (participation)
>
> I pretended to participate in the activities (distraction)
>
> I rushed through the activities (distraction)

In addition to these changes, while we found that the reliability for the evaluation factor was strong, we chose to add a third item to strengthen the factor: “I would recommend this instructor to other students” (Cashin, 1988, 1990). As our objective to this study was to measure the effects of in-class exercises, we also modified or removed all instances of out-of-class learning from our instrument to represent only those types of instruction that occur during class. Following this modification process, we finalized the StRIP instrument v1.0 (Appendix). This instrument represents our team’s efforts to further investigate students’ responses to different types of instruction and is ready to be administered as part of our full-scale study.

(1)Limitations and Future Research

It is worth noting a few limitations in our instrument development, which we plan to address in our future research. First, the exploratory and confirmatory factor analyses are based on data from eight courses at four institutions. Although the four institutions represent doctoral, baccalaureate, and minority-serving institutions, our findings are not necessarily generalizable. Furthermore, although our sample sizes appear to meet recommendations from the literature, they are still small and might influence our model fit. In our future research, we plan to expand our data collection methods to more locations to address these issues. We also plan to refine the instrument as needed on the basis of these expanded results.

Second, because we asked students how often they reacted in various ways to all activities as a whole, rather than specific types of activities (see Appendix), it is more difficult to relate specific activities to specific student reactions. This decision was a tradeoff for brevity, because students would have to respond to the 17 student response to instruction items for each of the 21 different instructional types listed. We may reconsider expanding this survey in future studies to focus on student response to specific types of instruction.

Third, the estimates for our types of instruction models are based solely on ideal types of instruction. As noted earlier in the Exploratory Factor Analysis section, we believe this arises because it would be difficult for an instructor to actually cover each of these types of instruction in a semester. However, students still perceive these types of instruction as aligned with one of these four categories: interactive, constructive, active, and passive. Consequently, much of our future research will directly investigate how students feel about these types of instruction in their ideal classroom, whether or not this perception aligns with what they actually experienced in the classroom, and subsequently, how they responded to the use of these types of instruction. Furthermore, we also plan to consider the use of separate constructs for the actual and ideal types of instruction in our future research.

Finally, the instrument relies on student self-reports of instructional practices, instructor strategies, and reactions to active learning. While this limitation is less of a concern for positivity, value, and evaluation items (Table 3), student reports of their own participation (Table 3), instructor strategies (Table 2), and frequency of types of instruction (Table 1) may be different from those of other students and the instructor. We will note this constraint in all our future research utilizing this instrument, yet we conclude that students' perception of the frequency of these activities is an important key to understanding how they ultimately respond to active learning. Other aspects of our ongoing work describe the preliminary results for our findings on student response to types of instruction and how we are comparing student and instructor responses and working with instructors to interpret their own data in instructional decisions.

(1)Conclusion

This article has described the design process and pilot results for an instrument to measure student response to instructional practices. Since our focus was on development of the instrument, future analyses will involve a broader administration and more systematic analysis of the instrument across multiple types of courses and institutions. The instrument measures three constructs related to our framework: types of instruction, strategies for using in-class activities, and student responses to instruction. Although the instrument was developed in the context of required gateway engineering courses, we expect that it may be relevant for a wider variety of STEM contexts, and we encourage other researchers to examine its usefulness in other contexts.

We believe there are several practical implications for the use of this instrument in the engineering classroom. First, we described a spectrum of activities in the instrument so instructors can examine the types of instruction currently used in the engineering classroom, and how ideal these activities might be to their students. Second, from our review of the literature, we compiled a list of several strategies for using in-class activities that instructors may wish to incorporate into their own courses to support student engagement. Third, we provided a list of students' responses to these types of instruction so that instructors can examine how their students respond to these activities and identify behaviors that might indicate students are disengaged during the process. Finally, our overall framework was developed with the hope that researchers and instructors, alike, can utilize this instrument to study multiple classrooms and identify relationships between types of instruction, how each type of instruction is introduced, and how students subsequently respond. For example, do students notice efforts taken by an instructor to explain the purpose of an activity? If not, maybe these efforts need to be more explicit or more frequent. Similarly, a few vocal students can sometimes give the impression that the entire class dislikes active learning. Having results from the instrument can help an instructor understand the views of all students in the class. There is much to be learned about this important area, and we encourage other instructors and researchers to use and build on this instrument in their own work.

// typesetter – OK to start Appendix on new page, if only small amount of text falls ono the bottom of last page of article //

Appendix

StRIP Student Instrument

Student Responses to Instruction[a]

In this course, when the instructor asked you to do an in-class activity (e.g., solve problems in a group during class or discuss concepts with classmates), how often did you react in the following ways?

I did not actually participate in the activity.
I gave the activity minimal effort.
I felt positively towards the instructor.
I tried my hardest to do a good job.
I distracted my peers during the activity.
I pretended to participate in the activity.
I felt the effort it took to do the activity was worthwhile.
I participated actively (or attempted to).
I talked with classmates about other topics besides the activity.
I felt the instructor had my best interests in mind.
I saw the value in the activity.
I felt the time used for the activity was beneficial.
I enjoyed the activity.
I surfed the internet, checked social media, or did something else

instead of doing the activity.

I rushed through the activity.

Strategies for Using In-Class Activities[a]

In this course, when the instructor asked you to do an in-class activity (e.g,. solve problems in a group during class or discuss concepts with classmates), how often did the instructor do the following things?

Clearly explained what I was expected to do for the activity.

Clearly explained the purpose of the activity.

Discussed how this activity related to my learning.

Solicited my feedback or that of other students about the activity.

Used activities that were the right difficulty level (not too easy, not too difficult).

Walked around the room to assist me or my group with the activity, if needed.

Encouraged students to engage with the activity through his/her demeanor.

Gave me an appropriate amount of time to engage with the activity.

Course Evaluation[b]

Please rate your level of agreement with the following items.

Overall, this was an excellent course.

Overall, the instructor was an excellent teacher.

I would recommend this instructor to other students.

Types of Instruction

For each of the following things, please indicate how often you did each thing in this course[c] and how often you would like to do each in your ideal course[d].

Listen to the instructor lecture during class.

Brainstorm different possible solutions to a given problem.

Find additional information not provided by the instructor to complete assignments.

Work in assigned groups to complete homework or other projects.

Make individual presentations to the class.

Be graded on my class participation.

Study course content with classmates outside of class.

Assume responsibility for learning material on my own.

Discuss concepts with classmates during class.

Make and justify assumptions when not enough information is provided.

Get most of the information needed to solve the homework directly from the instructor.

Be graded based on the performance of my group.

Preview concepts before class by reading, watching videos, etc.

Solve problems in a group during class.

Solve problems individually during class.

Answer questions posed by the instructor during class.

Ask the instructor questions during class.

Take initiative for identifying what I need to know.

Watch the instructor demonstrate how to solve problems.

Solve problems that have more than one correct answer.

Do hands-on group activities during class.

---

[a] Response options for each item were: 1 = almost never (<10% of the time); 2 = seldom (~30% of the time); 3 = sometimes (~50 % of the time); 4 = often (~70 % of the time); 5 = very often (>90 % of the time). [b] Response options for each item were: 1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree. [c] Response options for each item were: 1 = never; 2 = seldom (1–5 times per semester); 3 = sometimes (5–10 times per semester); 4 = often (once a week); 5 = very often (more than once/week). [d] Response options for each item were: 1 = much less; 2 = slightly less; 3 = about the same; 4 = slightly more; 5 = much more.

Acknowledgments

This project was funded by the U.S. National Science Foundation through grants 1347417, 1347482, 1347580, 1347718, and 1500309. The opinions are those of the authors and do not necessarily represent the National Science Foundation. The authors would like to thank the instructors and students who agreed to be part of the pilot study, as well as project advisory board members.

References

American Association for the Advancement of Science (AAAS). (2010). *Vision and change: A call to action.* Washington, DC: AAAS. http://visionandchange.org/files/2011/03/VC-Brochure-V6-3.pdf

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME). (2014). *Standard for educational and psychological testing*. Washington, DC: AERA.

Angelo, T. A., & Cross. K. P. (1993). *Classroom assessment techniques. A handbook for college teachers.* San Francisco, CA: Jossey-Bass.

Armstrong, J. S. (1998). Are student ratings of instruction useful? *American Psychologist, 53*(11), 1223–1232. http://dx.doi.org/10.1037/0003-066X.53.11.1223

Arum, R., & Roksa, J. (2011). *Academically adrift.* Chicago, IL: University of Chicago Press. http://dx.doi.org/10.7208/chicago/9780226028576.001.0001

Appleton, J. J., Christenson, S. L., & Furlong, M. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in Schools, 45*, 369–386. http://dx.doi.org/10.1002/pits.20303

Bacon, D., Stewart, K., & Silver, W. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Mangement Education, 23*(5), 467–488. http://dx.doi.org/10.1177/105256299902300503

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science*, *16*(1), 74–94. http://dx.doi.org/10.1007/BF02723327

Bentley, F. J., Kennedy, S., & Semsar, K. (2011). How not to lose your students with concept maps. *Journal of College Science Teaching, 41*, 61–68. http://www.colorado.edu/sei/documents/IPHY/HowNotToLoseStudentsWithConceptMaps_JCST2011.pdf

Borrego, M., Froyd, J. E., & Hall, T. S. (2010). Diffusion of engineering education innovations: A survey of awareness and adoption rates in US engineering departments. *Journal of*

Engineering Education, 99(3), 185–207. http://dx.doi.org/10.1002/j.2168-9830.2010.tb01056.x

Brophy, J. E. (1987). Socializing students' motivation to learn. In M. L. Maehr & D. Kleiber (Eds.), *Advances in motivation and achievement: Enhancing motivation* (pp. 181–210). Greenwich, CT: JAI Press.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming.* New York: Routledge

Carberry, A. R., Lee, H. S., & Ohland, M. W. (2010). Measuring engineering design self-efficacy. *Journal of Engineering Education, 99*(1), 71–79. http://dx.doi.org/10.1002/j.2168-9830.2010.tb01043.x

Cashin, W. E. (1988). Student ratings of teaching: A summary of the research. *IDEA Paper No. 20*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, W. E. (1990). Student ratings of teaching: Recommendations for use. *IDEA Paper No. 22*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

Chasteen, S. (2014, November 4). Measuring and improving students' engagement [Blog post]. Retrieved January 9, 2016 from http://blog.sciencegeekgirl.com/2014/11/02/measuring-and-improving-students-engagement/.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*, 73–105. http://dx.doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. http://dx.doi.org/10.1080/00461520.2014.965823

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.]

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. Gunnar & L. A. Sroufe (Eds.), *Minnesota Symposium on Child Psychology* (Vol. 23). Chicago, IL: University of Chicago Press.

Dancy, M. H., & Henderson, C. (2012). Experiences of new faculty implementing research-based instructional strategies. *Proceedngs of the 2011 Physics Education Research Conference*, Melville, NY. http://homepages.wmich.edu/~chenders/Publications/2011DancyPERCPaper.pdf

Donohue, S., & Richards, L. (2009). *Factors affecting student attitudes toward active learning activities in a graduate engineering statistics course*. Paper presented at the 39th ASEE/IEEE Frontiers in Education conference, San Antonio, TX. http://dx.doi.org/10.1109/FIE.2009.5350587

Eisinga, R., Grotenhuis, M. T., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?. *International Journal of Public Health, 58*(4), 637–642. http://dx.doi.org/10.1007/s00038-012-0416-3

Felder, R. M. (2011). Hang in there! Dealing with student resistance to learner-centered teaching. *Chemical Engineering Education, 43*(2), 131–132. http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Columns/HangInThere.pdf

Felder, R. M., & Brent, R. (2009). Active learning: An introduction. *ASQ Higher Education Brief, 2*(4), 1–5. http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Papers/ALpaper(ASQ).pdf

Finelli, C. J., Daly, S. R., & Richardson, K. M. (2014). Bridging the research-to-practice gap: Designing an institutional change plan using local evidence. *Journal of Engineering Education, 103*(2), 331–361. http://dx.doi.org/10.1002/jee.20042

Finn, J. D., Folger, J., & Cox, D. (1991). Measuring participation among elementary school students. *Educational and Psychological Measurement, 50*, 393–401. http://dx.doi.org/10.1177/0013164491512013

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. http://dx.doi.org/10.2307/3151312

Fraser, B. J. (1998). Classroom environment instruments: Development, validity and applications. *Learning Environments Research, 1*, 7–33. http://dx.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. *Review of Educational Research, 74*(1), 59–109. http://dx.doi.org/10.3102/00346543074001059

Friedrich, K., Sellers, S., & Burstyn, J. (2007). Thawing the chilly climate: Inclusive teaching resources for science, technology, engineering, and math. *To Improve the Academy: Resources for Faculty, Instructional, and Organizational Development, 26*, 133–144.

Froyd, J., Borrego, M., Cutler, S., Prince, M., & Henderson, C. (2013). Estimates of use of research-based instructional strategies in core electrical or computer engineering courses. *IEEE Transactions on Education, 56*(4), 393–399. http://dx.doi.org/10.1109/TE.2013.2244602

Furlong, M. J., & Christenson, S. L. (2008). Engaging students with school and academics: A relevant construct for all students. *Psychology in Schools, 45*, 365–368. doi:10.1002/pits.20303

Gaffney, J. D. H., Gaffney, A. L. H., & Beichner, R. J. (2010). Do they see it coming? Using expectancy violation to gauge the success of pedagogical reforms. *Physical Review Special Topics – Physics Education Research, 6*(1). http://dx.doi.org/10.1103/PhysRevSTPER.6.010102

Gauci, S. A., Dantas, A. M., Williams, D. A., & Kemm, R. E. (2009). Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiology Education*, *33*(1), 60–71. http://dx.doi.org/10.1152/advan.00109.2007

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational evaluation and policy analysis*, *11*(3), 255–274. http://dx.doi.org/10.2307/1163620

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis.* New York, NY:Macmillan.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., . . . Wood, W. B. (2004). Scientific teaching. *Science, 304*(5670), 521–522. http://handelsmanlab.sites.yale.edu/sites/default/files/Scientific%20Teaching.pdf

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238–247. http://dx.doi.org/10.1037/1040-3590.7.3.238

Henderson, C., & Dancy, M. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics - Physics Education Research, 3*(2), 020102-020101 to 020102-020114. http://dx.doi.org/10.1103/PhysRevSTPER.3.020102

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104–121. http://dx.doi.org/10.1177/109442819800100106

Hora, M. T., Ferrare, J., & Oleson, A. (2012). *Findings from Classroom Observations of 58 Math and Science Faculty*. Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research. http://ccher.wceruw.org/documents/CCHER_final%20report_2012.pdf

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. http://dx.doi.org/10.1080/10705519909540118

Hung, W., Bailey, J. H., & Jonassen, D. H. (2003). Exploring the tensions of problem-based learning: Insights from research. *New Directions for Teaching and Learning, 2003*(95), 13–23. http://dx.doi.org/10.1002/tl.108

Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a Job in General scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology*, *74*(2), 193. http://dx.doi.org/10.1037/0021-9010.74.2.193

Jamieson, L. H., & Lohmann, J. R. (2012). *Innovation with impact: Creating a culture for scholarly and systematic innovation in engineering education*. Washington, DC: American Society for Engineering Education. https://www.asee.org/member-resources/reports/Innovation-with-Impact/Innovation-With-Impact-Report.pdf

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). *Active learning: Cooperation in the college classroom*. Edina, MN: Interaction.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*, 187–200. http://dx.doi.org/10.1007/BF02289233

Kline, R.B. (1998). *Principles and practice of structural equation modeling.* New York, NY: The Guilford Press.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges and Universities. http://provost.tufts.edu/celt/files/High-Impact-Ed-Practices1.pdf

Lake, D. (2001). Student performance and perceptions of a lecture-based course compared with the same course utilizing group discussion. *Physical Therapy, 81*, 896–902.

Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research*, *83*(3), 432–479. http://dx.doi.org/10.3102/00346vv54313480891

Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education*, *68*, 241–270. http://dx.doi.org/10.2307/2112741

Li, Q., McCoach, D. B., Swaminathan, H., & Tang, J. (2008). Development of an instrument to measure perspectives of engineering education among college students. *Journal of Engineering Education, 97*(1), 47–56. http://dx.doi.org/10.1002/j.2168-9830.2008.tb00953.x

Livingstone, D., & Lynch, K. (2000). Group project work and student-centred active learning: Two different experiences. *Studies in Higher education*, *25*(3), 325–345. http://dx.doi.org/10.1080/713696161

Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, *33*(3), 269–298. http://dx.doi.org/10.1177/0011000004272260

Martens, M. P., & Hasse, R. F. (2006). Advanced applications of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, *34*, 878–911. http://dx.doi.org/10.1177/0011000005283395

Michael, J. (2007). Faculty perceptions about barriers to active learning. *College Teaching, 54*(1), 42–47. http://dx.doi.org/10.3200/CTCH.55.2.42-47

Moffett, B. S., & Hill, K. B. (1997). The transition to active learning: A lived experience. *Nurse Educator, 22*(4), 44–47. http://dx.doi.org/10.1097/00006223-199707000-00015

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine (NAS, NAE, & IOM). (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC. https://www.utsystem.edu/competitive/files/RAGS-fullreport.pdf

Newmann, F., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11–39). New York, NY: Teachers College Press.

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*, 248–292.

Oakley, B. A., Hanna, D. M., Kuzmyn, A., & Felder, R. M. (2007). Best practices involving teamwork in the classroom: Results from a survey of 6435 engineering student respondents. *IEEE Transactions on Education, 50*(3), 266–272. http://dx.doi.org/10.1109/TE.2007.901982

Ouimet, J. A., Bunnage, J. C., Carini, R. M., Kuh, G. D., & Kennedy, J. (2004). Using focus groups, expert advice, and cognitive interviews to establish the validity of a college student survey. *Research in Higher Education, 45*(3), 233–250. http://dx.doi.org/10.1023/B:RIHE.0000019588.05470.78

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*(1), 115–135.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist, 37*(2), 91–105.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-2018-7_12

Pintrich, P. R., & De Groot, E. (1990). Motivated and self-regulated learning components of academic performance. *Journal of Educational Psychology*, *82*, 33–40. http://rhartshorne.com/fall-2012/eme6507-rh/cdisturco/eme6507-eportfolio/documents/pintrich%20and%20degroodt%201990.pdf

Powell, K. C., & Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education*, *130*(2), 241.

President's Council of Advisors on Science and Technology (PCAST). (2012). *Report to the president: Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: Executive Office of the President. https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-executive-report-final_2-13-12.pdf

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–232. http://dx.doi.org/10.1002/j.2168-9830.2004.tb00809.x

Prince, M., Borrego, M., Henderson, C., Cutler, S., & Froyd, J. (2013). Use of research-based instructional strategies in core chemical engineering courses. *Chemical Engineering Education*, *47*(1), 27–37. http://homepages.wmich.edu/~chenders/Publications/2013PrinceCEEUseofRBIS.pdf

Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education*, *95*(2), 123–138. http://dx.doi.org/10.1002/j.2168-9830.2006.tb00884.x

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287–297. http://dx.doi.org/10.1037/1040-3590.12.3.287

Ro, H. K., Merson, D., Lattuca, L. R., & Terenzini, P. T. (2015). Validity of the contextual competence scale for engineering students. *Journal of Engineering Education*, *104*(1), 35–54. http://dx.doi.org/10.1002/jee.20062

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592. http://dx.doi.org/10.1093/biomet/63.3.581

Seidel, S. B., & Tanner, K. D. (2013). What if students revolt? –Considering student resistance: Origins, options, and opportunities for investigation. *CBE-Life Sciences Education*, *12*(4), 586–595. http://dx.doi.org/10.1187/cbe-13-09-0190

Seymour, E., & Hewitt, N. (1997). *Talking about leaving*. Boulder, CO: Westview Press.

Shekhar, P., DeMonbrun, M., Borrego, M., Finelli, C., Prince, M., Henderson, C., & Waters, C. (2015). Development of an observation protocol to study undergraduate engineering student resistance to active learning. *International Journal of Engineering Education, 31*(2), 597–609. http://homepages.wmich.edu/~chenders/Publications/2015ShekharIJEE.pdf

Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (Eds.). (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press. http://eduinfo.cchem.berkeley.edu/pdf/NAPunderstandingandimprovinglearning.pdf

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effect of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*, 571–581. http://dx.doi.org/10.1037/0022-0663.85.4.571

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21–44). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-2018-7_2

Stipek, D. (2002). Good instruction is motivating. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (pp. 309–332). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-012750053-9/50014-0

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York, NY: Harper & Row.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling a typology with examples. *Journal of mixed methods research, 1*(1), 77–100. http://dx.doi.org/10.1177/2345678906292430

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10694-000

Van Barneveld, A., & Strobel, J. (2011). *Reports from teaching practice: Experiences and management of tensions encountered with PBL implementations in the early years of undergraduate engineering education*. Paper presented at the Research in Engineering Education Symposium, Madrid, Spain.

Wehlage, G. G., Rutter, R. A., Smith, G. A., Lesko, N. L., & Fernandez, R. R. (1989). *Reducing the risk: Schools as communities of support.* Philadelphia: Farmer Press.

Weimer, M. (2002). *Learner-centered teaching: Five key changes to practice*. San Francisco, CA: Jossey-Bass.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781412983655

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research a content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806–838. http://dx.doi.org/10.1177/0011000006288127

Yadav, A., Subedi, D., Lunderberg, M., & Bunting, C. (2011). Problem-based Learning: Influence on Students' Learning in an Electrical Engineering Course. *Journal of Engineering Education, 100*, 253–280. http://dx.doi.org/10.1002/j.2168-9830.2011.tb00013.x

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 21*, 3–17. http://dx.doi.org/10.1207/s15326985ep2501_2

(1)Authors

Matt DeMonbrun is a doctoral candidate in the School of Education at the University of Michigan, 610 East University, Ann Arbor, MI, 48109; mdemonbr@umich.edu.

Cynthia J. Finelli is an associate professor of electrical engineering and computer science, associate professor of education, and Director of the Engineering Education Research Program at the University of Michigan, 4227 EECS Building, 1301 Beal Avenue, Ann Arbor, MI, 48109; cfinelli@umich.edu.

Michael Prince is a professor of chemical engineering at Bucknell University, Lewisburg, PA, 17837; prince@bucknell.edu.

Maura Borrego is an associate professor of mechanical engineering and STEM education at the University of Texas at Austin, 204 E. Dean Keeton Street, Stop C2200, ETC II 5.160, Austin, TX, 78712; maura.borrego@austin.utexas.edu.

Prateek Shekhar is a post-doctoral research fellow in the department of biomedical engineering at the University of Michigan, 1101 Beal Avenue, Ann Arbor, MI, 48109; pshekhar@umich.edu.

Charles Henderson is a professor of physics at Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, MI, 49008; charles.henderson@wmich.edu.

Cindy Waters is an assistant professor of mechanical engineering at North Carolina A&T State University, 1601 E. Market Street, McNair Hall Room 614, Greensboro, NC, 27411; kwaters@ncat.edu.