Research in Higher Education
https://doi.org/10.1007/s11162-021-09625-z

How Students' Intellectual Orientations and Cognitive Reasoning Abilities and May Shape Their Perceptions of Good Teaching Practices

K. C. Culver[1] · **Nicholas A. Bowman**[2] · **Ernest T. Pascarella**[2]

Received: 13 February 2020 / Accepted: 25 January 2021

Abstract
Recent research has uncovered significant concerns about the validity of some types of college student self-reports. This study examines the extent to which student reports about a critical type of college experience—good teaching practices—may be biased as a function of students' intellectual orientations and cognitive reasoning abilities. Perceptions of instruction and instructional practices are especially important in higher education, given their increasing use for institutional quality assurance, as well as faculty rehiring and promotion processes. Using a large, multi-institutional, longitudinal dataset of first-year students, this study shows that several cognitive indicators predict perceptions of six different sets of good teaching practices and that these relationships do not seem to be explained by actual differences in students' experiences. Additional analyses indicate that halo effects, in which global evaluations of instructor quality and institutional satisfaction affect students' perceptions of their engagement with good practices, may partially explain these findings. The results provide important implications for practice and research related to college student survey data, including ways that these biases can be reduced or eliminated to more accurately capture students' engagement in good practices and the factors that may contribute to students' perceptions of their environment.

Keywords Student self-reports · Student ratings · Good teaching practices · Cognitive traits · Survey bias

A considerable body of research exists on instructional practices in higher education; effective pedagogical techniques have been shown to promote students' content learning, cognition, psychosocial development, and success in college (see Pascarella and Terenzini 1991, 2005; Perry and Smart 2007; Mayhew et al. 2016). National longitudinal assessments of college students, such as the National Survey of Student Engagement (NSSE), the Cooperative Institutional Research Program (CIRP), and the Wabash National Study of Liberal

✉ K. C. Culver
kculver@usc.edu

[1] Pullias Center for Higher Education, University of Southern California, Los Angeles, CA, USA

[2] The Center for Research on Undergraduate Education, University of Iowa, Iowa City, IA, USA

Arts Education (WNS), aim to measure students' progress on these and other liberal outcomes articulated by institutions and policymakers (e.g., Boyer Commission 1998; American Association of Colleges and Universities 2011) while also gathering vast data about students' engagement in the institution. Although a great deal of higher education research uses data from these and similar datasets, past researchers have examined the limitations of college student surveys in terms of construct and content validity (Campbell and Cabrera 2011; Porter 2011); in other words, these scholars question how well surveys accurately measure the concepts of interest. There have also been more specific examinations of students' ability to self-report gains (Bowman 2010a, b) and to self-report the impact of college experiences (Bowman and Seifert 2011). However, little research has examined how students' abilities and orientations may affect their survey responses. If these characteristics contribute to students' responses in systematic ways, understanding these relationships can allow researchers to account for these biases and therefore better estimate the impact of students' college experiences on their outcomes.

The present paper explores this issue as it relates to students' perceptions of effective teaching. In synthesizing the existing evidence on instructional practices, Chickering and Gamson (1987) created seven broad principles for good instructional practice in undergraduate education: (a) encouraging student–faculty contact, (b) encouraging cooperation among students, (c) encouraging active learning, (d) giving prompt feedback to students, (e) emphasizing time on task, (f) communicating high expectations, and (g) respecting diverse talents and ways of learning. Since their inception, these seven principles have been vetted by the literature (Chickering and Gamson 1999; Sorcinelli 1991), and a wealth of empirical studies have demonstrated their validity for predicting outcomes (for a review, see Pascarella et al. 2006). Recent longitudinal, multi-institutional studies have found widespread benefits of exposure to good teaching practices on students' cognitive and affective growth, educational persistence, and career outcomes (e.g., Cruce et al. 2006; Jessup-Anger 2012; Padgett et al. 2009; Pascarella et al. 2010; Padgett 2011; Pascarella et al. 2011; Loes et al. 2014). Because of the benefits of effective instruction for students' learning, development, and success, students' self-reported exposure to these practices are used for accountability both at the institutional level and at the level of individual instructors (Campbell and Cabrera 2011). Thus, the accuracy of scores specific to instructional practices have implications that are far-reaching, as these scores are increasingly used for faculty rehiring and promotion processes and for accreditation. Internationally, measures of teaching quality are also used in performance funding assessments (De Boer et al. 2015).

To contextualize this issue, we begin by providing an overview of psychological frameworks of survey response. We then discuss the conditions under which college student surveys may be more (or less) valid, and we move toward contextualizing these issues through theory and scholarship on teaching and learning and students' ratings of instruction. Finally, we argue that two categories of cognitive characteristics likely influence students' self-reported perceptions of instructional practices: students' cognitive reasoning abilities (including skills such as judgment and evaluation) and their intellectual orientation to instructional environments (including their motivations and dispositions to learn and think critically).

Theory and Research on Survey Self-Reports

Perhaps the most frequently used model of the psychology of survey response was proposed by Tourangeau et al. (2000). Drawing upon considerable evidence, they assert that survey respondents may engage in four sequential processes when answering a specific question: (1) *comprehension* of what the question is asking, (2) *retrieval* of relevant information, (3) *judgment* of the completeness and relevance of memories, and (4) choosing a *response* from among the available options. The accuracy of responses generally depends on participants' willingness and ability to engage sufficiently in each of the steps, along with the cognitive demands required to provide a correct answer.

College student surveys often include questions that vary substantially in their difficulty. On one side of the spectrum, students are asked personal information about their social identities and academic achievement for which a clear, correct answer exists and that answer is known (or knowable) to the respondent, making the cognitive tasks of retrieval and judgment relatively simple. As one example, Kuncel et al. (2016) conducted a meta-analysis of the validity of self-reported grades and SAT test scores; they found that the correlations between self-reported values and institutional records were quite high for grades ($r = 0.84$) and for total SAT score ($r = 0.82$). However, some systematic patterns in misreporting were apparent. First, students with high grades and test scores were much more accurate than those with low grades and test scores; this trend may occur because higher-performing students are more motivated to know their exact scores, since they may be more likely to apply to selective undergraduate and graduate programs as well as merit-based scholarships. Second, most (but not all) of the inaccurate self-reports were higher than the actual values, which suggests that social desirability, or the tendency to respond in a way that is perceived to be more socially acceptable, explains at least some of the inaccuracy of students' responses.

On the other side of the difficulty spectrum, students are often asked to report how much they have learned or changed while attending college. Such questions pose substantial challenges for students, since these questions may include vague language to describe the outcomes of interest, making it complicated in terms of comprehension; students may have limited information with which to make this assessment; students may have difficulty judging which memories are relevant, and students may have difficulty mapping their self-report onto a response option (they may also be influenced by social desirability to say that they have gained a great deal). As a result, students' responses to college self-reported gain questions are subject to various systematic biases that scholars posit stem from students' flawed retrieval and judgment processes, as well as the social desirability of responses (Pike 1993, 1999; Ross 1989; Bowman and Brandenberger 2010; Bowman 2011a; Bowman and Hill 2011; Porter 2013). Further, students' self-reported gains are only weakly correlated with longitudinal assessments that purportedly measure the same outcome (Gosen and Washbush 1999; Hess and Smythe 2001; Bowman 2010a, b, 2011b; Bowman and Brandenberger 2010).

That said, most items on college student surveys fall somewhere between these two extremes. Students are frequently asked about their engagement in a variety of college experiences within and outside of coursework. These questions typically provide response options that contain either specific frequencies that are potentially knowable but difficult to estimate (e.g., 11–15 h/week) or vague options that assess subjective impressions of frequency (e.g., "often" or "rarely"). Each approach seems better suited to a particular type of engagement. For instance, the number of hours that students spend in the classroom or

studying outside of class can both be estimated, whereas their instructors' use of active learning strategies or clarity of instruction does not fit neatly into this form of time-based measurement.

Considerable debate exists around the quality of college student survey responses more generally. Porter (2011) offers a substantial critique and ultimately questions whether current college student surveys provide any useful information. Barge and Gehlbach (2012) found evidence of widespread survey "satisficing," in which college students took excessive shortcuts to minimize the amount of time and effort in responding (e.g., choosing the same response every time; quitting the survey early). However, McCormick and McClenney (2012) have argued that many of the critiques of the validity of college student surveys are narrow or unfounded. In addition, Chen (2011) found that only a small proportion of participants (less than 10%) provide low-quality survey responses, arguing that these participants can be identified and removed to provide more accurate results. At the same time, it is certainly possible that individual differences can result in differences in perceptions of good teaching; in fact, two students enrolled in the same course section could respond differently to the instruction given (Pascarella 2001).

Conceptual Framework

This study expands upon previous research to examine how measures of cognitive reasoning abilities and intellectual orientations may affect students' perceptions of their instructors' use of good teaching practices. We ground the study in Entwistle's (2010) model of the interacting influences on student learning, adding to it theory and research specific to survey measurement. Based on decades of research on postsecondary teaching and learning, Entwistle's model identifies two main categories of influences on student learning: features of the teaching–learning environment and the characteristics of students. Students' perceptions of teaching are at the center of his framework, at the intersection of their individual characteristics and the teaching–learning environment.

Entwistle (2010) positions instructional practices as central to the teaching–learning environment. He also specifies a number of factors that influence these practices, including structural aspects of the course, instructors' beliefs about the role of the teacher, and the "inner logic" (p. 30) of the subject being taught. For instance, previous studies have suggested that instructors' use of good practices varies according to their disciplinary affiliation (Braxton et al. 1998; Kilgo et al. 2017).

The model also includes two categories of student characteristics that influence the ways they approach learning (Entwistle 2010). The first category includes relatively stable traits such as intelligence, profile of abilities, skills, and learning approaches. The second category includes students' orientations and dispositions, including their motives, feelings, and organized effort, which may be more easily shaped by students' classroom experiences. Further, the model acknowledges the role of students' backgrounds and previous experiences by articulating that students' family members, friends, previous instructors, and other mentors influence both categories of the influences on their learning.

Previous scholars have demonstrated the role that students' approaches to learning play in shaping their perceptions of teaching. For instance, students who implement deep approaches to learning by analyzing and synthesizing information are more likely to have a positive view of teaching practices that promote conceptual understanding, such as those outlined by Chickering and Gamson (1987); in contrast, students using surface approaches

to learning, acquiring content for course success rather than for understanding have a more negative view of the same practices (Entwistle and Tait 1990; Prosser and Trigwell 1990).

The role of students' cognitive characteristics has also been explored in terms of survey measurement specific to students' ratings of instructions (SRI; also known as course evaluations or students' evaluations of teaching). In particular, much of the variance in students' responses to SRI can be explained by four factors: students' level of interest in the subject matter, the number of students taking the course as an elective rather a requirement, their expected grades, and their perceptions of higher course workload/difficulty (for a review, see Marsh 2007). Another study found that students' "cognitive style" (p. 640), as measured by the year of enrollment and major field, shapes their responses, such that courses are rated more favorably by advanced undergraduates who have developed more active and complex approaches to thinking and when the course discipline is aligned with students' preferred ways of thinking (Ting 2000).

Another source of cognitive bias explored in the literature on survey measurement is specific to evaluations of other people (Tourangeau et al. 2000). In early research on performance evaluation, Thorndike (1920) identified the halo/horns effect, where a respondent's impression of one critical aspect of another person leads to a more global positive or negative evaluation, which in turn influences their perceptions and specific ratings more widely. In a couple of classic examples of the halo/horns effect, college student ratings of the pleasantness of an instructor's accent were notably affected by whether the instructor was presented as nice or as unfriendly (Nisbett and Wilson 1977), and even the perceived usefulness of a course textbook differed dramatically depending on whether the instructor was engaging versus boring in class (Williams and Ceci 1997).

Present Study and Hypotheses

Entwistle's (2010) model, along with the research on which it based, provide the foundation for our supposition that students' cognitive traits may influence their self-reported exposure to good teaching practices. Additionally, scholarship on SRI suggests that individual instructors and courses are rated more favorably when students are motivated to learn and willing to think deeply about course content, when they are cognitively challenged by course difficulty, and when the ways of thinking practiced in a course align with their own thinking skills and preferences. Whereas this scholarship is based on students' ratings specific to one instructor/course, college student surveys generally ask students to report their experiences in courses and with instructors at an aggregate level. In fact, we uncovered no research that explored the role of students' cognitive reasoning abilities and their orientations to learning with a focus on surveys of students' college experiences.

Therefore, to separately examine each of the two categories of students' characteristics outlined by Entwistle (2010), this study examines the relationship of students' cognitive reasoning abilities and their intellectual orientations with their perceptions of good teaching practices in their first year of college. We used a large, multi-institutional sample from one of the most comprehensive college student studies conducted in the last twenty years. Based on existing literature and our conceptual framework, we expected that students would interpret their instructors' practices through the lens of their own abilities and motivations, which would then result in different assessments of these practices. Specifically, our hypotheses are as follows:

1. Students who are highly motivated and have a strong inclination toward lifelong learning will be more attuned to the positive instructional practices that they receive. Given these students' greater receptivity toward learning, they should be likely to interpret the same behaviors as being more effective, and their motivation may also make them more likely to notice instructors' behaviors.
2. Conversely, students with greater cognitive reasoning abilities will be less likely to perceive good instructional practices. Such students may be more critical when considering many domains of life (including instructors' behaviors), and they may feel that established standards that are not particularly challenging to achieve. Further, these students may use a greater threshold for the meaningfulness of instructional engagement and the frequency of that engagement.
3. The association between cognitive measures and students' perceptions of instructional practices will be at least partially explained by the halo/horns effect. Specifically, accounting for students' global perceptions of instruction and satisfaction will reduce or eliminate these observed relationships.

While the use of a multi-institutional dataset allowed us to test these hypotheses in a college student survey, we also wanted to ensure that, to the extent possible, we were able to distinguish differences in students' reports of instruction among students who received similar or identical instruction. Therefore, our analyses used institutional fixed effects to account for between-college variation. At the student level, we incorporated a wide variety of control variables, including demographics, high school academic engagement, college academic engagement (e.g., honors, research), undergraduate major, paradigmatic development of the discipline in students' coursework, and college grades. The analyses also focused on the first year of college when students primarily take general education courses and therefore experience less variation in instructional practices within and across disciplines. Providing an even more stringent test, additional analyses examined subgroups of students who are likely to have had rather substantial overlap in their first-year coursework.

The focus on perceptions of instruction is a critical feature of this study, since such measures are used for high-stakes decisions, including accreditation and performance funding. Moreover, six different indicators of instruction were created to reflect various domains of Chickering and Gamson's (1987) seven principles for good practice. The number and diversity of instructional practice measures, along with the consideration of four cognitive indicators, also helped address competing explanations for the findings. For instance, motivated students might engage in objectively greater levels of student-faculty contact outside of the classroom, whereas it is unlikely that this motivation could lead to more prompt feedback. It also seems quite unlikely that students with greater cognitive reasoning ability would have actual lower objective exposure to good instructional practices.

Methods

Sample

The data for this paper come from the Wabash National Study of Liberal Arts Education (WNS). Funded by the Center of Inquiry in the Liberal Arts at Wabash College, the WNS is a longitudinal, pre-test/post-test study examining the cognitive and affective student

learning outcomes focusing on experiences theoretically associated with liberal arts education (see King et al. 2007). The WNS includes three cohorts of students (entering college in 2006, 2007, and 2008) from 48 institutions that varied by institutional characteristics, including type and control, selectivity, size, geographic location, and mission. The individuals in the sample were first-year, full-time undergraduate students enrolled at participating institutions; at smaller institutions, the initial sample included the entire incoming first-year class, while participants were selected randomly from the incoming first-year class at large institutions.

The data utilized in our study were collected at two timepoints. At the beginning of the fall semester of their first year of college, students completed a questionnaire that asked about background characteristics, including demographic information, family background, high school experiences, educational and career aspirations, and political and religious orientations. Students also completed previously vetted instruments measuring attitudes and dispositions, including the Academic Motivation scale and the Need for Cognition scale, which measures how much people enjoy engaging in effortful cognitive activities, among other instruments. In addition, students were randomly selected to complete either the Collegiate Assessment of Academic Proficiency (CAAP) Critical Thinking Test or the Defining Issues Test 2, because of the time involved in completing the instruments,. The second data collection occurred in the following spring. In addition to completing the same cognitive and psychosocial instruments assessed upon entering college, students completed the National Survey of Student Engagement (NSSE) and the WNS Student Experiences Survey. These instruments include items about students' experiences, levels of engagement, and exposure to good teaching practices.

The analytic sample for this study was 6,531 students. Because only half of students in the study completed CAAP, the analytic sample for this outcome was 3,004 students. Within the full sample and the CAAP subsample, the proportions of students by race and sex were consistent, as 8 percent of participants were Black/African American, 5 percent were Asian/Pacific Islander, 5 percent were Latino/Hispanic, 3 percent were another race/ethnicity (or race/ethnicity was unknown), and 38 percent were male. Descriptive statistics for the sample and each measure included in this study are provided in Appendix A.

Dependent Variables

The dependent variables measure students' perceptions of the extent to which their faculty members used six of Chickering and Gamson's (1987) good practices; WNS data does not include items that reliably measure the seventh good practice (respecting diverse talents and ways of learning). Student-faculty contact outside the classroom was indicated with a 9-item scale ($\alpha = 0.77$) that combined items measuring (a) students' perceptions of faculty members' interest and willingness to interact with students outside the classroom, and (b) the frequency with which students interacted with faculty members outside the classroom. Active learning was assessed via a 10-item scale ($\alpha = 0.77$) regarding students' perceptions of the frequency that faculty members used various active learning techniques (including class discussions, class presentations, and assignments that required critique of an argument). Collaborative learning was measured with a 7-item scale ($\alpha = 0.71$) about students' perceptions of the frequency that their instructors integrated collaborative learning approaches inside and outside of class (i.e., group projects, study groups). Prompt feedback was indicated with a 3-item scale ($\alpha = 0.67$) regarding students' perceptions that instructors provided timely written or oral feedback and evaluated student learning informally in the

classroom. Time on task was examined via an 8-item scale (α = 0.72) combining measures of (a) the amount of effort students put forth on studying as a result of institutional and instructor expectations, and (b) students' perceptions that instruction was relevant, organized, and helpful for achieving clearly defined course goals. Finally, high expectations were assessed with a 5-item scale (α = 0.81) about students' perceptions of the frequency that instructors challenged students intellectually, especially through the use of techniques requiring higher-order thinking (i.e., applying, critiquing, and/or arguing).

Good practice scales have been empirically vetted through use in previous studies of classroom experiences in undergraduate education (Pascarella et al. 2004, 2005, 2006; Cruce et al. 2006; Seifert et al. 2014). This study modifies existing good practice scales to avoid items that are largely dependent on student behaviors and attitudes (e.g., "Extent that my non-classroom interactions with faculty have had a positive influence on my intellectual growth and interest in ideas") or that require a high degree of subjective judgment beyond perceptions (e.g., "Frequency that faculty made good use of examples and illustrations to explain difficult points"). The items included in each scale are listed in Appendix B. Each dependent variable measures students' perceptions of the use of that good teaching practice in their first-year classes.

Key Independent Variables

Our study has four independent variables of interest, grouped into two categories reflective of Entwistle's (2010) model: cognitive reasoning abilities and intellectual orientation. In the cognitive reasoning abilities category, we examined students' ACT composite score (or SAT equivalent) and critical thinking skills. The latter outcome was measured using the CAAP critical thinking module, which is a 40-min, 32-item instrument designed to measure a student's ability to clarify, analyze, evaluate, and extend arguments. The internal consistency reliability for the CAAP ranges between 0.81 and 0.82 (ACT 1991).

Two distinct variables composed the intellectual orientation category: academic motivation and need for cognition. Academic motivation was measured using an eight-item scale (α = 0.69). Inclination to inquire and lifelong learn was measured through the Need for Cognition Scale (NCS). Need for cognition refers to an individual's "tendency to engage in and enjoy effortful cognitive activity" (Cacioppo et al. 1996, p. 197). The NCS is measured through an 18-item scale with high internal reliability (α = 0.89). Students who score high on the NCS scale tend to be lifelong learners who engage in inquiry, thinking, and reflection to make sense of their world, while students who score low on the NCS scale are more likely to rely on rules and the advice or actions of others. Each of the independent variables of interest was measured in the first data collection timepoint, which occurred before students had exposure to the instructional practices used in their courses.

Two additional variables were used to explore potential mechanisms that might indicate the presence of the halo/horns effect. Perceptions of instructor quality was measured through an index of two items (α = 0.89): "Most faculty with whom I have had contact are outstanding teachers" and "Most faculty with whom I have had contact are genuinely interested in teaching" (1 = strongly disagree, to 5 = strongly agree). As a more global measure of college quality, we also included an overall college satisfaction scale (α = 0.72) computed as the index of two items: "How would you evaluate your entire educational experience at this institution?" (1 = poor, to 4 = excellent), and "If you could start over again, would you go to the same institution you are now attending?" (1 = definitely no, to 4 = definitely yes).

Control Variables

Numerous control variables were included to account for the external influences on students and the teaching–learning environment outlined by Entwistle (2010). Demographic variables included race (dummy variables for Black/African American, Asian/Pacific Islander, Latinx/Hispanic, and other/unknown, with White/Caucasian as the reference group), sex (0 = female, 1 = male), and first-generation college student (0 = no, 1 = yes). To account for students' proclivity for engaging with their instructors, the frequency of high school interactions with teachers outside class was also used (1 = never, to 5 = very often).

Several different approaches to measuring students' coursework were employed. First, as a general indicator, students' undergraduate major was reported through a series of dummy variables (biological science, business, education, engineering, physical science, professional, social science, other, and undecided, with arts and humanities as the referent group). Second, drawing upon institutional records of specific courses that each student took during the first year, we created a scale of paradigmatic development based on Biglan's (1973) categorization of hard and soft disciplines to measure students' unique course-taking patterns. In hard disciplines such as physics and engineering, the high degree of consensus about standards, methods, and processes influences pedagogical norms, such that instructors in these fields are less likely to use some good practices than are instructors in soft fields such as business and sociology (Braxton et al. 1998; Kilgo et al. 2017). Third, given substantial differences in classroom practices between courses that do and do not cover diversity-related content (e.g., Nelson Laird and Engberg 2011), three variables indicated the frequency of taking each of the following course types (0 = 0 courses, to 4 = 4 or more courses): diverse cultures and perspectives (e.g., ethnic studies), women's/gender studies, and those that focus on issues of equality and/or social justice. Fourth, engagement in several specific course or program types was also indicated with separate variables (0 = no, 1 = yes): honors program or college, first-year seminar, learning community, service learning, and undergraduate research.

Two additional student-level indicators were included. College grades have a consistently strong relationship with student ratings of instruction (for a review, see De Witte and Rogge 2011), and grades could reflect different levels of engagement in certain course practices, so these were used (1 = C- or lower, to 8 = A). Hours working for pay may decrease the amount of time that students can dedicate to their coursework, so this was also included (1 = 0 h, to 8 = more than 30 hours).

Analysis

To make the sample more representative of first-year students at these institutions, we developed and implemented a multi-level sample weighting algorithm for use in analyses (for more information about the construction and application of sampling weights, see Rabe-Hesketh and Skrondal 2006; Groves et al. 2009). The institution-level weight accounts for random sampling at larger institutions. The individual-level weight incorporates race, sex, and standardized test scores to adjust for non-response bias, as groups that tend to have lower rates of college persistence were also more likely to drop out of the sample over time (i.e., males, Blacks/African Americans, Latinos/Hispanics, and students with lower standardized test scores).

To examine the extent to which students' intellectual orientations and cognitive abilities predict their reports of good teaching, we conducted a series of multiple regression analyses. Institutional fixed effects were used; this approach is ideal for solely examining within-institution dynamics (Allison 2009). Institutional fixed effects accounts for the variation between institutions, including unobserved characteristics of each institution like tenure and promotion policies that might shape students' exposure to good teaching practices. Our preferred model specification is given by Eq. (1):

$$Y_{ij} = \beta_1 cognitive_i + \beta X_i + \mathrm{T}campus_j + \varepsilon_i$$

In this equation, our outcome Y_{ij} is the reported exposure to one of the six good practice variables for student i attending institution j. Our independent variable of interest $cognitive_i$ is one of the four precollege indicators of students' cognitive traits, and X_i is a vector of student-level covariates (e.g., demographics, college coursework). As we implemented institutional fixed effects, $campus_j$ is a vector of dummy variables indicating all but one of the institutions, and ε_i is the error term. Each precollege indicator of students' intellectual orientation and cognitive reasoning ability measure was included in a separate model to reduce multicollinearity, so 24 total analyses were conducted within the full sample. All continuous variables were standardized with a mean of zero and a standard deviation of one to facilitate the interpretation and comparison of unstandardized regression coefficients (Cohen et al. 2003; Mayhew et al. 2016).

In an effort to compare students whose curricular experiences and exposure to instructional practices were as similar as possible, we conducted the same analyses within two key subsets of our sample: students in honors programs and students who declared an engineering major. Honors programs and honors colleges generally use a cohort model, in which students are taking many classes together by definition (Slavin et al. 2008; Ogilvie and Reza 2009); in fact, a study of honors colleges found that 97 percent offered honors courses to fulfill general education requirements (Sederberg 2005). Honors students also tend to be highly motivated and high achieving academically as a necessary condition for admittance, thereby reducing the range on both types of cognitive measures in this study. Another group of students who take similar coursework are those who major in engineering. The course requirements for engineering are quite prescribed by accreditation, so students within the same institution—and often across institutions—must take very similar courses in their first year before they enter their specialty (e.g., electrical, mechanical, civic; see ABET 2016). As a result, limiting the analytic sample to two different subgroups further reduces the possibility that the observed relationships are the product of actual student differences in experiencing good teaching practices.

We then conducted additional analyses to investigate whether any link between these precollege cognitive measures and instructional perceptions was the product of the halo/horns effect, which is the tendency for a global positive or negative evaluation of a person's performance to influence specific ratings and perceptions. These analyses incorporated two forms of global evaluation as additional predictors: overall perceptions of instructor quality and overall college satisfaction. If halo error can account for the observed relationships, then these global perceptions should explain the link between the cognitive measure and perceptions of good practices.

We examined our models for potential multicollinearity by computing a correlation matrix and variance inflation factor tests (VIFs). The correlation matrix showed no correlations over $r = 0.50$. The variance inflation factors (VIFs) for variables of interest in overall models ranged from 1.17 to 1.67, while the VIFs of students' cognitive traits ranged from

1.19 to 2.67 in moderation and subgroup analyses, suggesting there were no significant collinearity concerns with the models (Allison 1999; Stevens 2002).

Limitations

This study includes some limitations. Institutions that participated in WNS were invited to participate based on a stated commitment to liberal arts education and location in the United States; therefore, these institutions may not be representative of all colleges and universities. As such, the results of this study may not necessarily be generalizable to all colleges and universities or all college students. Additionally, this study uses an existing dataset where measures of intellectual ability and intellectual orientation were determined by researchers who designed WNSLAE, which limits this study's interpretation to these specific measures. Finally, this study seeks to identify bias in measures of good teaching, but we have no way of determining students' actual objective engagement with these practices, which would be helpful for unequivocally demonstrating bias in student self-reports. Instead, we have taken a different approach that involves ruling out alternative hypotheses for our findings, as described in detail below.

Results and Discussion

Overall, intellectual orientation and cognitive reasoning abilities are consistently related to perceptions of good practices in the expected direction. The results for the primary analyses within the full sample are presented in Table 1; these employ institutional fixed effects and control for student demographics, high school involvement, undergraduate major, various coursework characteristics, and college grades. Academic motivation is positively and significantly related to each of the six measures of good teaching practices (βs $\geq$ 0.14, ps $<$ 0.001), with standardized coefficients considered medium to large effect sizes according to guidelines for college impact research (Mayhew et al. 2016). Need for cognition is also positively and significantly associated with all self-reports of good teaching practices (βs $\geq$ 0.09, ps $<$ 0.001), reflecting medium effect sizes. These findings demonstrate the expected pattern, such that students who are interested in and excited about cognitively challenging academic work report higher levels of exposure to good instructional practices. Importantly, we believe that some—if not all—of this observed relationship is an artifact of students' perceptions that are independent of their actual level of exposure. Although it seems possible that more academically motivated students could have greater student-faculty interaction even in the presence of substantial statistical controls, do motivated students actually receive feedback more promptly than their counterparts? One would have to make this questionable assertion to explain the significant findings observed here.

Also as expected, critical thinking scores are inversely and significantly related to student-faculty interaction, active learning, collaborative learning, and time on task (βs $\leq$ -0.06, ps $<$ 0.01). ACT composite scores, which serve as another indicator of cognitive reasoning ability, are also negatively associated with these same measures of good teaching (βs $\leq$ -0.11, ps $<$ 0.001). We believe that these negative relationships reflect the more critical lens through which these students may view their classroom experiences. These relationships are smaller in magnitude compared to those found for intellectual orientations, reflecting small to medium effect sizes. At the same time, findings are impressive when

Table 1 Institutional fixed effects analyses examining the relationship between students' cognitive reasoning abilities/orientations and their perceptions of instructors' use of good practices

	Academic motivation	Need for cognition	Critical thinking	ACT ability
	β	β	β	β
Good practice outcome	(SE)	(SE)	(SE)	(SE)
Student-faculty interaction	0.21***	0.16***	−0.08***	−0.11***
	(0.01)	(0.01)	(0.02)	(0.02)
Active learning	0.18***	0.14***	−0.07**	−0.13***
	(0.01)	(0.01)	(0.02)	(0.02)
Collaborative learning	0.14***	0.09***	−0.06**	−0.11***
	(0.01)	(0.01)	(0.02)	(0.02)
Prompt feedback	0.14***	0.13***	0.04	0.04
	(0.01)	(0.01)	(0.02)	(0.02)
Time on task	0.17***	0.14***	−0.10***	−0.12***
	(0.01)	(0.02)	(0.02)	(0.02)
High expectations	0.18***	0.13***	0.00	−0.01
	(0.01)	(0.01)	(0.02)	(0.02)
N	6489	6487	3044	6531

All continuous variables were standardized. Institutional fixed effects models were used that included controls for students' race/ethnicity, sex, first-generation status, high school interactions with teachers, intended major, paradigmatic development of college coursework, hours spent working for pay, and participation in the following academic experiences in college: first-year seminars, learning communities, service learning, undergraduate research, diversity courses, women's/gender studies courses, and courses on equality/social justice. Each cognitive measure was included in a separate model to reduce multicollinearity

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

considering that the included measures of students' intellectual orientation and cognitive reasoning ability are positively correlated with one another, but the relationships of these characteristics with good teaching measures diverge in theoretically expected directions.

When examining results specific to students in honors programs, the vast majority of the significant relationships from the full-sample analyses persist within this subsample (see Table 2). In particular, academic motivation is positively related to all good practice measures (βs ≥ 0.08, ps < 0.01), and need for cognition is positively related to five of six good practices (βs ≥ 0.08, ps < 0.05). The significant relationships for ACT composite scores are also replicated (βs ≤ -0.11, ps < 0.01), and a previously non-significant link with instructors' high expectations becomes significant in this subsample ($\beta = -0.11$, $p < 0.05$). The standardized coefficients for these three measures predominantly reflect medium to large effect sizes. The lone difference among this subsample is for critical thinking skills, as this indicator of cognitive reasoning ability only significant predicts time on task among honors students ($\beta = -0.15$, $p < 0.05$). However, it is important to note that only half of the sample completed the critical thinking test as part of the WNS research design. Therefore, the number of honors students within this analytic sample is fairly modest (N = 486), especially when considering the inclusion of 30 control variables and use of fixed effects for 46 institutions. This lack of statistical power likely explains why regression coefficients of

Table 2 Institutional fixed effects analyses among honors students and engineering students examining the relationship between students' cognitive reasoning abilities/orientations and their perceptions of instructors' use of good practices

	Honors students				Engineering students			
	Academic motivation	Need for cognition	Critical thinking	ACT ability	Academic motivation	Need for cognition	Critical thinking	ACT ability
	β	β	β	β	β	β	β	β
Good practice outcome	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)
Student-faculty interaction	0.25***	0.18***	−0.04	−0.16***	0.33***	0.30***	−0.01	−0.18
	(0.03)	(0.04)	(0.06)	(0.05)	(0.06)	(0.08)	(0.10)	(0.11)
Active learning	0.13***	0.11**	−0.12	−0.21***	0.18**	0.17*	0.00	−0.25*
	(0.03)	(0.03)	(0.06)	(0.05)	(0.07)	(0.07)	(0.09)	(0.10)
Collaborative learning	0.08*	0.04	−0.10	−0.15**	0.28***	0.17*	−0.01	−0.25*
	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.08)	(0.11)	(0.11)
Prompt feedback	0.17***	0.14***	0.05	0.01	0.21**	0.26***	0.14	0.13
	(0.03)	(0.03)	(0.06)	(0.05)	(0.07)	(0.08)	(0.12)	(0.11)
Time on task	0.18***	0.16***	−0.15*	−0.16**	0.12	0.15*	0.08	−0.03
	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.08)	(0.10)	(0.11)
High expectations	0.15***	0.10**	−0.07	−0.11*	0.29***	0.24**	0.05	−0.02
	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.08)	(0.11)	(0.10)
N	1061	1058	486	1066	296	297	166	296

All continuous variables were standardized. Institutional fixed effects models were used that included controls for students' race/ethnicity, sex, first-generation status, high school interactions with teachers, intended major, paradigmatic development of college coursework, hours spent working for pay, and participation in the following academic experiences in college: first-year seminars, learning communities, service learning, undergraduate research, diversity courses, women's/gender studies courses, and courses on equality/social justice. Each cognitive measure was included in a separate model to reduce multicollinearity

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

−0.10 and −0.12 were not significant for critical thinking, but the same effect size was significant for other key predictors. Taken as a whole, these subsample results provide further evidence that student self-reports of good teaching practices may be influenced by their intellectual orientations and cognitive abilities.

We also examined these relationships among engineering students, as students within the same institution likely experienced a relatively prescribed first-year curriculum. At the same time, a fairly small number of students had declared engineering majors upon entering the first year (N = 295), so statistical power is a concern. That said, the analyses within this subsample identified numerous significant results (see Table 2). Five of six good practice outcomes are significantly predicted by academic motivation (βs ≥ 0.18, ps < 0.01) and all six outcomes are predicted by need for cognition (βs ≥ 0.15, ps < 0.05). ACT composite score is also inversely and significantly related to active learning and collaborative learning (βs = −0.25, ps < 0.05). The effect sizes are even higher here than those in the full-sample analyses for which the possibility of students taking substantively different courses is a more salient concern. Among this subgroup, several of the standardized coefficients are a quarter of a standard deviation or more, which is notable given the number of controls included in the models and the examination of students who take a common set of courses. Critical thinking is not significantly associated with any good practice among engineering students, but the very small sample size (N = 166) likely played a role in limiting these results.

The results for analyses examining the potential role of the halo/horns effect appear in Table 3. The pattern of significant results is identical to that for the primary analyses shown in Table 1. The strength of the relationships declines within some of the analyses containing academic motivation (βs ≥ 0.09, ps < 0.001) and need for cognition (βs ≥ 0.06, ps < 0.001), which suggests that students' global perceptions of the quality of their instructors and institutional satisfaction explain some of the link between intellectual orientation and perceptions of exposure to good practices. In contrast, the inclusion of these global measures does not diminish the magnitude of relationships for the cognitive reasoning ability measures of critical thinking (βs ≤ −0.06, ps < 0.01) and ACT composite (βs ≤ -0.11, ps < 0.001); if anything, these results exhibit very slight increases over those from the original analyses (although these would certainly not differ significantly). Further, perceptions of instructor quality (βs = 0.15–0.42, ps < 0.001) and institutional satisfaction (βs = 0.14–0.22, ps < 0.001) are both significantly related to all six good practice measures. Students' reports of the presence of good practices may therefore be partially explained by their positive evaluation of their instructors and institutions; perhaps not surprisingly, instructor quality perceptions appear to be (modestly) more useful at explaining this link than are global perceptions of institutional satisfaction. This study cannot conclusively determine why this examination of halo error explains some of the results for intellectual orientation, but not for cognitive reasoning ability. That said, these findings are interesting to consider in light of Entwistle's (2010) argument that students' dispositions and motivation are more pliable than their abilities and knowledge. Because students' instructional experiences directly shape their dispositions, it may be that students who have intellectual orientations are more likely to have a generally favorable view of their instructors. We posit that this halo effect extends even beyond instructors, manifesting as a love of learning, which is why it explains part of the relationship of intellectual orientations with perceptions of practices.

Table 3 Institutional fixed effects analyses including controls for perceptions of instructors' quality and institutional satisfaction examining the relationship between students' cognitive reasoning abilities/orientations and their perceptions of instructors' use of good practices

	Academic motivation	Need for cognition	Critical thinking	ACT ability
Good practice outcome	B (SE)	B (SE)	B (SE)	B (SE)
Student-faculty interaction	0.15***	0.09***	−0.09***	−0.12***
	(0.01)	(0.01)	(0.02)	(0.02)
Active learning	0.15***	0.11***	−0.08***	−0.13***
	(0.01)	(0.01)	(0.02)	(0.02)
Collaborative learning	0.11***	0.06***	−0.07**	−0.11***
	(0.01)	(0.01)	(0.02)	(0.02)
Prompt feedback	0.09***	0.08***	0.03	0.03
	(0.01)	(0.01)	(0.02)	(0.02)
Time on task	0.11***	0.08***	−0.11***	−0.12***
	(0.01)	(0.01)	(0.02)	(0.02)
High expectations	0.14***	0.08***	−0.00	−0.01
	(0.01)	(0.01)	(0.02)	(0.02)
N	6489	6487	3044	6531

All continuous variables were standardized. Institutional fixed effects models were used that included controls for students' demographics (students' race/ethnicity, sex, and first-generation status), high school and college engagement in academic experiences (high school interactions with teachers, intended major, paradigmatic development of college coursework, hours spent working for pay, perceptions of instructor quality, and participation in the following academic experiences in college: first-year seminars, learning communities, service learning, and undergraduate research), college first-year academic curriculum (intended major, paradigmatic development of courses, diversity courses, women's/gender studies courses, and courses on equality/social justice), hours spent working for pay, instructional quality, and institutional satisfaction. Each cognitive measure was included in a separate model to reduce multicollinearity

*$p < 0.05$

**$p < 0.01$

***$p < 0.001$

Conclusions and Implications

This paper provides intriguing evidence that students' perceptions of exposure to good teaching may be informed by their intellectual orientation and cognitive reasoning ability. These results persist when incorporating a wide array of relevant control variables, when accounting for all differences across institutions through fixed effects, and when limiting the sample to honors or engineering students who are likely to be taking many courses together. The medium to large effects sizes found for intellectual orientation appear to be partially, but not entirely, explained by the halo/horns effect. Relationships for cognitive reasoning abilities were comparatively smaller in magnitude, but the pattern of results was consistent. In short, student reports of exposure to good practices may be biased in predictable ways by students' cognitive traits.

That said, it is critical to clarify some key features of our argument, conclusions, and implications. First, student self-reports are likely the best way to collect large-scale data on students' curricular experiences, especially when considering how these experiences are associated with their individual learning and success outcomes. These self-reports still

likely reflect actual exposure to good practices to a large extent; this view is supported by studies linking such measures to a variety of desired college outcomes (for reviews, see Pascarella and Terenzini 2005; Mayhew et al. 2016). Some research has illustrated the possibility of measuring certain aspects of instructor behavior through direct classroom observations using trained raters (see Campbell 2017), and this approach shows considerable promise. However, it may present some challenges in terms of scalability for understanding the experiences of many students.

Second, the fact that these potential biases are predictable helps considerably with overcoming them when conducting research. While several frameworks of college impact name cognitive traits as importance sources of variation among students (e.g., Pascarella 1985; Astin 1970; Braxton et al. 2004; Museus 2014), characteristics such as demographics and high school academic achievement are more frequently used. Thus, if scholars and institutional researchers are to obtain an accurate estimate of the unique effect of students' perceptions of instruction on their learning and development, they will likely need to take students' intellectual abilities and orientations into account in addition to the other factors that have been conceptually and empirically linked to differences in the ways that students engage and develop in higher education. Otherwise, results will probably include confounded and inflated estimates of students' perceptions of instruction. Controlling for students' ACT scores and academic motivation in multivariate analyses may largely, if not entirely, remove the biases we identified in students' reports of teaching practices. This usage is particularly important when drawing comparisons across groups, whether these pertain to coursework taught by different instructors, in different fields of study, or at different institutions. Accounting for these student inputs reduces the likelihood that any observed differences can be attributed to students' motivation and their global perceptions rather than actual instructional or institutional practices. Considerations such as these may prove especially beneficial in international contexts, given increasing attention to measuring teaching quality as part of global rankings and to determine institutional funding (De Boer et al. 2015; Altbach and Hazelkorn 2018).

Third, intellectual orientations and cognitive abilities are not the only attributes that may affect the validity of student reports of instructional practice. As noted earlier, students' perceptions of academic experiences may be the product of various factors, including students' expected course grades (Spooren et al. 2013). This study included a host of covariates that may also shape students' perceptions of college experiences, and these should also be incorporated into future assessment and research. Conducting comparative analyses with and without these covariates may also yield insights into the extent to which such factors affect the results and corresponding conclusions.

Fourth, intellectual orientation and cognitive reasoning ability are not necessarily static over time or across contexts. We used measures of each of these constructs that were obtained at a particular point in time; we chose the beginning of the first year, since these may then serve as a lens through which students viewed their coursework over the next two semesters. However, these measures all changed to some extent over the course of a year and even more so over four years (among students who participated in all three waves of the WNS). The use of the term "ability," then, does not imply a fixed mindset or a lack of malleability. In addition, although we operationalized intellectual orientation and cognitive reasoning ability broadly, these constructs may vary by course and by subject matter. Many students will clearly be interested in the content of some courses more than others; in addition, students who have strong writing skills but weaker quantitative skills may perceive an English course quite differently than they would perceive a physics course. Future research

might explore the degree to which these systematic biases persist among fourth-year students, especially in terms of whether students may become less biased raters over time.

Moreover, future research should not only integrate the practical implications of this study (by accounting for academic motivation and cognitive reasoning ability when studying good practices) but also further investigate issues pertaining to the effective measurement of students' exposure to good teaching practices within and across institutions. In particular, the dynamics examined in this study should also be explored at the individual course level in terms of student ratings of instruction. Such examination is especially pertinent as student course evaluations are used not only for individual instructional improvement, but also for individual appraisal and institutional accountability (Spooren et al. 2013). Students' ratings of instruction are a primary means through which college instructors are evaluated for their teaching performance, which then affects decisions about their continued employment, raises, and promotions; additionally, these evaluations are often included as part of institutional assessment for accreditation. As such, examining the potential relationships of students' cognitive traits with their perceptions of individual instructors could have important implications for college faculty and academic administrators. For instance, are students' orientations and abilities that are most relevant to a particular course more important than the domain-general attributes that we examined here? Are these apparent biases less problematic for "low-inference" instructor behaviors that are easily observable and measurable (e.g., specific attributes of course assignments) versus "high-inference" behaviors that are more opaque (e.g., overall clarity and organization of instruction)? Such questions will help in the effort to understand the best approaches for examining college teaching and desired outcomes.

Appendix A Variable descriptions, descriptive statistics, and reliabilities for ACT sample (N = 6531).

Variable	Definition	Mean	Standard deviation	Min	Max
Intellectual orientations and cognitive reasoning abilities					
Precollege academic motivation	Mean-based scale measuring academic motivation; 8-item scale, α = 69; standardized	0.00	1.00	−4.47	2.50
Precollege need for cognition	The degree to which one enjoys engaging in effortful cognitive activities; 18-item scale, α = 0.89; standardized	0.00	1.00	−3.82	2.50
Precollege ability (ACT or equivalent)	Composite ACT or SAT equivalent score converted to an ACT metric standardized	0.00	1.00	−3.77	2.22
Precollege critical thinking	Critical thinking skills; 32-item scale, α = 0.81—0.82; standardized	0.00	1.00	−2.85	1.94

Variable	Definition	Mean	Standard deviation	Min	Max
Good teaching practices					
Student-faculty interactions	Students' perceptions of faculty's interest and willingness to interact with students outside the classroom and the frequency with which students interacted with faculty outside the classroom; 9-item scale, α = 0.77; standardized	0.00	1.00	−3.67	2.85
Active learning	Students' perceptions of the frequency that faculty used various active learning techniques (including class discussions, class presentations, and assignments that required critique of an argument); 10-item scale, α = 0.77; standardized	0.00	1.00	−3.63	2.34
Collaborative learning	Students' perceptions of the frequency that faculty integrated collaborative learning approaches inside and outside of class (i.e., group projects, study groups); 7-item scale, α = 0.71; standardized	0.00	1.00	−3.50	2.63
Prompt feedback	Students' perceptions that faculty provided timely written or oral feedback and evaluated student learning informally in the classroom; 3-item scale, α = 0.67; standardized	0.00	1.00	−3.00	2.27
Time on task	Students' perceptions that instruction was relevant, organized, and helpful for achieving clearly defined course goals and the amount of effort students put forth on studying as a result of institutional and instructor expectations; 8-item scale, α = 0.72; standardized	0.00	1.00	−5.91	2.49
High expectations	Students' perceptions of the frequency that faculty challenged students intellectually, especially through the use of techniques requiring higher-order thinking (i.e., applying, critiquing, and/or arguing); 5-item scale, α = 0.81; standardized	0.00	1.00	−3.36	2.00
Potential mechanisms					
Perceptions of instructor quality	Generalized assessment of teaching ability and interest in teaching among instructors; 2-item scale, α = 0.89; standardized	0.00	1.00	−3.94	1.41

Variable	Definition	Mean	Standard deviation	Min	Max
Overall college satisfaction	Overall satisfaction with educational experience at this institution; 2-item scale, α = 0.72; standardized	0.00	1.00	−3.68	1.10
Student background characteristics					
Race/Ethnicity: Black; African American	0 = no; 1 = yes	0.09	0.28	0.00	1.00
Race/Ethnicity: Asian; Pacific Islander	0 = no; 1 = yes	0.05	0.23	0.00	1.00
Race/Ethnicity: Latinx	0 = no; 1 = yes	0.05	0.21	0.00	1.00
Race/Ethnicity: Other; Race/Ethnicity Unknown	0 = no; 1 = yes	0.03	0.18	0.00	1.00
Sex: Male	0 = no; 1 = yes	0.38	0.48	0.00	1.00
First-Generation Student (no parent attended college)	0 = no; 1 = yes	0.11	0.31	0.00	1.00
H.S. Interactions with Teachers	Frequency of interacting with teachers outside of class during high school; (1 = never, to 5 = very often); standardized	0.00	1.00	−2.06	0.49
First-year college experiences					
Paradigmatic development of courses taken		0.00	1.00	−1.48	2.68
Intended major: biological sciences	0 = no; 1 = yes	0.11	0.31	0.00	1.00
Intended major: business	0 = no; 1 = yes	0.11	0.32	0.00	1.00
Intended major: education	0 = no; 1 = yes	0.07	0.26	0.00	1.00
Intended major: engineering	0 = no; 1 = yes	0.05	0.21	0.00	1.00
Intended major: physical science	0 = no; 1 = yes	0.06	0.24	0.00	1.00
Intended major: professional	0 = no; 1 = yes	0.11	0.32	0.00	1.00
Intended major: social science	0 = no; 1 = yes	0.17	0.37	0.00	1.00
Courses: diverse cultures and perspectives (e.g., ethnic studies)	Number of courses taken focusing on diverse cultures and perspectives (1 = 0 courses, to 5 = 4 or more courses)	0.00	1.00	−0.71	3.67
Courses: women's/gender studies	Number of courses taken focusing on women's/gender studies (1 = 0 courses, to 5 = 4 or more courses)	0.00	1.00	−0.42	5.86
Courses: focus on issues of equality and/or social justice	Number of courses taken focusing on issues of equality and/or social justice (1 = 0 courses, to 5 = 4 or more courses)	0.00	1.00	−0.67	4.14
Honors program or college	0 = no; 1 = yes	0.16	0.37	0.00	1.00
First-year seminar	0 = no; 1 = yes	0.67	0.47	0.00	1.00
Learning community	0 = no; 1 = yes	0.32	0.46	0.00	1.00
Service learning	0 = no; 1 = yes	0.45	0.59	0.00	1.00
Undergraduate research	0 = no; 1 = yes	0.05	0.22	0.00	1.00
College grades	Students' self-reported grades (1 = C- or lower, to 8 = A)	0.00	1.00	−3.24	1.22

Variable	Definition	Mean	Standard deviation	Min	Max
Hours working for pay	How many hours per week students spend working for pay (1 = 0 h to 8 = more than 30 h)	0.00	1.00	−0.68	8.12

Appendix B Items included in each scale of good teaching practices

Faculty-student interactions ($a = 0.76$)
Most faculty with whom I have had contact are genuinely interested in students[a]
Most faculty with whom I have had contact are willing to spend time outside of class to discuss issues of interest and importance to students[a]
Most faculty with whom I have had contact are interested in helping students grow in more than just academic areas[a]
During current school year, how often have you discussed grades or assignments with an instructor?[b]
During current school year, how often have you talked about career plans with a faculty member or advisor?[b]
During current school year, how often have you discussed ideas from readings or classes with faculty members outside of class?[b]
During current school year, how often have you worked with faculty members on activities other than coursework (committees, orientation, student life activities, etc.)?[b]
Indicate the extent to which you agree/disagree that I am satisfied with the opportunities to meet and interact informally with faculty members[a]
Collaborative learning (a = 0.70)
In my classes, students taught each other in addition to faculty teaching[c]
Faculty encouraged me to participate in study groups outside of class[c]
I have participated in one or more study group(s) outside of class. [c]
In your experience at your institution during the current school year, about how often have you worked with classmates outside of class to prepare class assignments?[b]
Active learning (a = 0.73)
In your experience at your institution during the current school year, about how often have you asked questions in class or contributed to class discussions?[b]
In your experience at your institution during the current school year, about how often have you made a class presentation?[b]
During the current school year, how much has your coursework emphasized synthesizing and organizing ideas, information, or experiences into new, more complex interpretations and relationships?[d]
During the current school year, how much has your coursework emphasized making judgments about the value of information, arguments, or methods, such as examining how others gathered and interpreted data and assessing the soundness of their conclusions?[d]
How often have exams or assignments required me to write essays[c]
How often have exams or assignments required me to use course content to address a problem not presented in the course[c]
How often have exams or assignments required me to compare or contrast topics or ideas from a course[c]

Active learning (a = 0.73)
How often have exams or assignments required me to point out the strengths and weaknesses of a particular argument or point of view[c]
How often have exams or assignments required me to argue for or against a particular point of view and defend my argument?[c]
Time on task (a = 0.72)
About how many hours in a typical week do you spend preparing for class (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities)?[f]
Frequency that faculty gave assignments that helped in learning the course material[c]
Frequency that class time was used effectively[c]
Frequency that the presentation of material was well organized[c]
Frequency that course goals and requirements were clearly[c] explained
High expectations (a = 0.71)
How often have faculty asked challenging questions in class?[c]
How often have faculty asked you to argue for or against a particular point of view?[c]
How often have faculty challenged your ideas in class?[c]
Mark the box that best represents the extent to which your examinations during the current school year challenged you to do your best work[e]
In your experience at your institution during the current school year, about how often have you worked harder than you thought you could to meet an instructor's standards or expectations?[b]
How often have students challenged each other's ideas in class?[c]
Prompt feedback (a = 0.67)
How often have faculty informed you of your level of performance in a timely manner?[c]
How often have faculty checked to see if you had learned the material well before going on to new material?[c]
In your experience at your institution during the current school year, about how often have you received prompt written or oral feedback from faculty on your academic performance?[b]

Response scales are as follows:

[a] 1 = strongly disagree to 5 = strongly agree

[b] 1 = never to 4 = very often

[c] 1 = never to 5 = very often

[d] 1 = very little to 4 = very much

[e] 1 = very little to 7 = very much

[f] 1 = 0 h to 8 = more than 30 hours

References

ABET. (2016). *Criteria for accrediting engineering programs, 2016–2017*. http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2016-2017/

Allison, P. D. (1999). *Multiple regression: A primer*. Pine Forge Press.

Allison, P. D. (2009). *Fixed effects regression models*. Sage.

Altbach, P. G., & Hazelkorn, E. (2018). Measuring education quality in global rankings: What's the likelihood? *International Higher Education, 95,* 12–14.

American Association of Colleges and Universities. (2011). *The LEAP vision for learning*. Washington, DC: American Association of Colleges and Universities.

American College Testing Program (ACT). (1991). *CAAP technical handbook*. Author.
Astin, A. W. (1970). The methodology of research on college impact, part one. *Sociology of Education, 43*(3), 223–254.
Barge, S., & Gehlbach, H. (2012). Using the theory of satisficing to evaluate the quality of survey data. *Research in Higher Education, 53*(2), 182–200.
Biglan, A. (1973). The characteristics of subject matter in difference academic areas. *Journal of Applied Psychology, 57*(3), 195–203.
Bowman, N. A. (2010). Can 1st-year college students accurately report their learning and development? *American Educational Research Journal, 47*(2), 466–496.
Bowman, N. A. (2010a). Assessing learning and development among diverse college students. In S. Herzog (Ed.), *Diversity and educational benefits* (New Directions for Institutional Research, no. 145, pp. 53–71). Jossey-Bass.
Bowman, N. A. (2011). Validity of self-reported gains at diverse institutions. *Educational Researcher, 40*(1), 22–24.
Bowman, N. A. (2011a). Examining systematic errors in predictors of college student self-reported gains. In S. Herzog & N. A. Bowman (Eds.), *Validity and limitations of college student self-report data* (New Directions for Institutional Research, no. 150, pp. 7–19). Jossey-Bass.
Bowman, N. A., & Brandenberger, J. W. (2010). Quantitative assessment of service-learning outcomes: Is self-reported change an adequate proxy for longitudinal change? In J. Keshen, B. Holland, & B. Moely (Eds.), *Research for what? Making engaged scholarship matter* (Advances in Service-Learning Research, Vol. 10, pp. 25–43). Information Age Publishing.
Bowman, N. A., & Hill, P. L. (2011). Measuring how college affects students: Social desirability and other potential biases in self-reported gains. In S. Herzog & N. A. Bowman (Eds.), *Validity and limitations of college student self-report data* (New Directions for Institutional Research, no. 150, pp. 73–85). Jossey-Bass.
Bowman, N. A., & Seifert, T. A. (2011). Can students accurately assess what affects their learning and development? *Journal of College Student Development, 52*(3), 270–290.
Boyer Commission on Educating Undergraduates in the Research University. (1998). *Reinventing undergraduate education: A blueprint for America's research universities.* The Carnegie Foundation for the Advancement of Teaching.
Braxton, J. M., Hirschy, A. S., & McClendon, S. A. (2004). *Understanding and reducing college student departure* (ASHE-ERIC Higher Education Report, Vol. 30, no. 3). Jossey-Bass.
Braxton, J. M., Olsen, D., & Simmons, A. (1998). Affinity disciplines and the use of principles of good practice for undergraduate education. *Research in Higher Education, 39*(3), 299–318.
Cacioppo, J., Petty, R., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253.
Campbell, C. M. (2017). An inside view: The utility of quantitative observation in understanding college educational experiences. *Journal of College Student Development, 58*(2), 290–299.
Campbell, C. M., & Cabrera, A. F. (2011). How sound is NSSE?: Investigating the psychometric properties of NSSE at a public, research-extensive institution. *The Review of Higher Education, 35*(1), 77–103.
Chen, P.-S.D. (2011). Finding quality responses: The problem of low-quality survey responses and its impact on accountability measures. *Research in Higher Education, 52*(7), 659–674.
Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39*(7), 3–7.
Chickering, A. W., & Gamson, Z. F. (1999). Development and adaptations of the seven principles for good practice in undergraduate education. *New Directions for Teaching and Learning*, *80*, 75–81.
Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum.
Cruce, T., Wolniak, G., Seifert, T., & Pascarella, E. (2006). Impacts of good practices on cognitive development, learning orientations, and graduate degree plans during the first year of college. *Journal of College Student Development, 47*(4), 365–383.
De Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., & Vossensteyn, H. (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Center for Higher Education Policy Studies.
De Witte, K., & Rogge, N. (2011). Accounting for exogenous influences in performance evaluations of teachers. *Economics of Education Review, 30*(4), 641–653.
Entwistle, N. (2010). *Taking stock: An overview of key research findings* (pp. 15–57). Taking stock: Research on teaching and learning in higher education.

Entwistle, N. J., & Tait, H. (1990). Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *Higher Education, 19,* 169–194.

Gosen, J., & Washbush, J. (1999). Perceptions of learning in TE simulations. *Developments in Business Simulation & Experiential Learning, 26,* 170–175.

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. John Wiley.

Hess, J. A., & Smythe, M. J. (2001). Is teacher immediacy actually related to student cognitive learning? *Communication Studies, 52*(3), 197–219.

Jessup-Anger, J. E. (2012). Examining how residential college environments inspire the life of the mind. *The Review of Higher Education, 35*(3), 431–462.

Kilgo, C. A., Culver, K. C., Young, R. L., & Paulsen, M. B. (2017). The relationship between students' perceptions of "good practices for undergraduate education" and the paradigmatic development of disciplines in course-taking behavior. *Research in Higher Education, 58*(4), 430–448.

King, P. M., Kendall Brown, M., Lindsay, N. K., & VanHecke, J. R. (2007). Liberal arts student learning outcomes: An integrated approach. *About Campus, 12*(4), 2–9.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2016). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63–82.

Loes, C. N., Salisbury, M. H., & Pascarella, E. T. (2014). Student perceptions of effective instruction and the development of critical thinking: A replication and extension. *Higher Education, 69*(5), 1–16.

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). Springer.

Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., & Wolniak, G. C., with Pascarella, E. T., & Terenzini, P. T. (2016). *How college affects students (Volume 3): 21st century evidence that higher education works*. Jossey-Bass.

McCormick, A. C., & McClenney, K. (2012). Will these trees *ever* bear fruit? A response to the special issue on student engagement. *The Review of Higher Education, 35*(2), 307–333.

Museus, S. D. (2014). The culturally engaging campus environments (CECE) model: A new theory of success among racially diverse college student populations. In *Higher education: Handbook of theory and research* (pp. 189–227). Springer.

Nelson Laird, T. F., & Engberg, M. E. (2011). Establishing differences between diversity requirements and other courses with varying degrees of diversity inclusivity. *The Journal of General Education, 60*(2), 117–137.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256.

Ogilvie, K., & Reza, E. M. (2009). Business student performance in traditional vs. honors course settings. *Business Education Innovation Journal, 1*(2), 31–37.

Padgett, R. D. (2011). *The effects of the first year of college on undergraduates' development of altruistic and socially responsible behavior*. (Doctoral dissertation). Iowa Research Online.

Padgett, R. D., Johnson, M. P., Saichaie, K., & Paulsen, M. B. (2009, November). *The impact of physical health and good practices on psychological well-being*. Paper presented at the annual conference of the Association for the Study of Higher Education, Vancouver, British Columbia, Canada.

Pascarella, E. T. (1985). College environmental influences on learning and cognitive development: A critical review and synthesis. *Higher Education: Handbook of Theory and Research, 1,* 1–61.

Pascarella, E. T. (2001). Using student self-reported gains to estimate college impact: A cautionary tale. *Journal of College Student Development, 42*(5), 488–492.

Pascarella, E. T., Cruce, T., Umbach, P., Wolniak, G., Kuh, G. D., Carini, R., et al. (2006). Institutional selectivity and good practices in undergraduate education: How strong is the link? *Journal of Higher Education, 77*(2), 251–285.

Pascarella, E. T., Cruce, T. M., Wolniak, G. C., & Blaich, C. F. (2004). Do liberal arts colleges really foster good practices in undergraduate education? *Journal of College Student Development, 45*(1), 57–74.

Pascarella, E. T., Salisbury, M. H., & Blaich, C. (2011). Exposure to effective instruction and college student persistence: A multi-institutional replication and extension. *Journal of College Student Development, 52*(1), 4–19.

Pascarella, E. T., Seifert, T. A., & Blaich, C. (2010). How effective are the NSSE benchmarks in predicting important educational outcomes? *Change, 42*(1), 16–22.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students*. San Francisco, CA: Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research (2nd Ed.)*. Jossey-Bass.

Pascarella, E., Wolniak, G., Seifert, T., Cruce, T., & Blaich, C. (2005). *Liberal arts colleges and liberal arts education: New evidence on impacts*. Jossey-Bass/ASHE.

Perry, R. P., & Smart, J. C. (2007). *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer.

Pike, G. R. (1993). The relationship between perceived learning and satisfaction with college: An alternative view. *Research in Higher Education, 34*(1), 23–40.

Pike, G. R. (1999). The constant error of the halo in educational outcomes research. *Research in Higher Education, 40*(1), 61–86.

Porter, S. R. (2011). Do college student surveys have any validity? *The Review of Higher Education, 35*(1), 45–76.

Porter, S. R. (2013). Self-reported learning gains: A theory and test of college student survey response. *Research in Higher Education, 54*(2), 201–226.

Prosser, M., & Trigwell, K. (1990). Student evaluations of teaching and courses: Student study strategies as a criterion of validity. *Higher Education, 20*(2), 135–142.

Rabe-Hesketh, S., & Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 169*(4), 805–827.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96*(2), 341–357.

Sederberg, P. (2005). Characteristics of the contemporary honors college. A descriptive analysis of a survey of NCHC member colleges. *Journal of the National Collegiate Honors Council, 6*(2), 121–136.

Seifert, T. A., Gillig, B., Hanson, J. M., Pascarella, E. T., & Blaich, C. F. (2014). The conditional nature of high impact/good practices on student learning outcomes. *The Journal of Higher Education, 85*(4), 531–564.

Slavin, C., Coladarci, T., & Pratt, P. A. (2008). Is student participation in a honors program related to retention and graduation rates? *Journal of the National Collegiate Honors Council, 9*(2), 59–69.

Sorcinelli, M. D. (1991). Research findings on the seven principles. In A. W. Chickering & Z. F. Gamson (Eds.), *Applying the seven principles of good practice in undergraduate education – New directions for teaching and learning* (No. 47) (pp. 13–25). Jossey Bass.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching the state of the art. *Review of Educational Research, 83*(4), 598–642.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences (4th ed.).* Erlbaum.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25–29.

Ting, K. F. (2000). A multilevel perspective on student ratings of instruction: Lessons from the Chinese experience. *Research in Higher Education, 41*(5), 637–661.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

Williams, W. M., & Ceci, S. J. (1997). "How'm I doing?" Problems with student ratings of instructors and courses. *Change: the magazine of higher learning, 29*(5), 12–23.