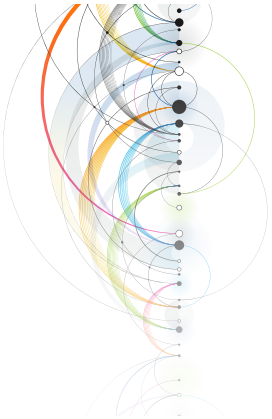

Abstract

Asking college students how much they have learned or grown is a common assessment practice in student affairs and elsewhere. Unfortunately, recent research suggests that these self-reported gains do a very poor job of measuring actual student learning and growth. This paper provides an overview of the psychological process of how students likely respond to such questions and why their responses can be seriously flawed. It also discusses circumstances in which self-reported gains are somewhat more valid and offers concrete suggestions for student affairs professionals and other higher education constituents who seek to accurately measure student outcomes.

AUTHOR

Nicholas A. Bowman, Ph.D.
Bowling Green State University

Understanding and Addressing the Challenges of Assessing College Student Growth in Student Affairs

In an era of increased demands for accountability and limited financial resources in higher education, the assessment of college student outcomes has become crucial. Several recent books have provided excellent guidelines and examples for conducting college student outcomes assessments (e.g., Astin & antonio, 2012; Banta, Jones, & Black, 2009; Suskie, 2009; Walvoord, 2010). In general, these authors agree that multiple forms of assessment should be administered, that direct assessments should be employed when possible, and that assessment results should inform programmatic and institutional change. To measure academic outcomes, many institutions are using standardized examinations (e.g., Collegiate Learning Assessment, Collegiate Assessment of Academic Proficiency) as well as "authentic assessments," such as portfolios or rubrics of student work (Kuh & Ikenberry, 2009). These indicators can be used to assess the achievement of a particular level of skill or competence and/or the amount of growth that has occurred during the undergraduate years.

However, such formalized, direct learning assessments are rarely used to measure the effectiveness of student affairs in promoting student outcomes. These rigorous assessments not only require a great deal of resources, but they also indicate the types of academic and general cognitive skills that are generally not considered to be the primary focus of student affairs. As a result, student affairs professionals use a variety of other approaches for measuring learning and growth, including responses to broad national surveys (e.g., National Survey of Student Engagement), specific national surveys (e.g., ACUHO-I/EBI Resident Assessment), and a variety of locally developed surveys (I recently heard about a written questionnaire assessing student experiences and outcomes from a residence hall ice cream social!). In many cases, outcomes assessment in student affairs simply involves asking students what they have learned and how they have grown. The responses to these questions are then interpreted as indicating students' actual learning and growth.

CORRESPONDENCE

Email
nbowman@bgsu.edu

In many cases, outcomes assessment in student affairs simply involves asking students what they have learned and how they have grown. The responses to these questions are then interpreted as indicating students' actual learning and growth.

Recent research has cast serious doubt upon the (seemingly reasonable) assumption that college students can accurately report their own growth. If these self-reports were accurate, then one would expect a high correlation between students' self-reported gains on a particular outcome (e.g., critical thinking skills) and longitudinal changes on a well-validated measure of that same outcome. Across various samples and outcomes, the correlations between longitudinal and self-reported gains on the same construct are consistently low ($rs < .20$), and they are often not significantly different from zero (Bowman, 2010a, 2010b, 2011b; Bowman & Brandenberger, 2010; Gosen & Washbush, 1999; Hess & Smythe, 2001). In addition, the significant predictors of longitudinal growth (e.g., college experiences, student demographics, institutional attributes) often diverge considerably from the significant predictors of self-reported gains for the same construct (Anaya, 1999; Bowman, 2010b, 2011a, 2012; Bowman & Brandenberger, 2010; Porter, 2013). As a result, practitioners and researchers would arrive at remarkably different conclusions about the experiences that promote or hinder student growth depending on the type of outcomes assessment that they use. Through a synthesis of the existing literature and examination of several theory-driven hypotheses, Porter (2013) argues that college self-reported gains should not be used as indicators of actual student learning. Finally, relevant to many student affairs assessments, college students also have considerable difficulty reporting the educational impact of a particular experience or set of experiences; in general, students tend to overestimate the effects that their experiences actually have (Bowman & Brandenberger, 2010; Bowman & Seifert, 2011; Conway & Ross, 1984).

In this paper, I will first discuss why students may have such a difficult time reporting their own growth and why their self-reports may not even reflect their actual judgments. Next, I will propose several conditions under which students provide somewhat more accurate assessments of their growth. Finally, I will provide suggestions for student affairs practitioners and other higher education constituents who seek to measure and understand student outcomes.

The Psychology of Student Self-Reported Gains

Practitioners and researchers would arrive at remarkably different conclusions about the experiences that promote or hinder student growth depending on the type of outcomes assessment that they use.

In their seminal work, Tourangeau, Rips, and Rasinski (2000) proposed a four-stage model of the psychology of survey responses. The four steps involved, in order, are comprehension of the question, retrieval of memories associated with the question, judgment of the completeness and relevance of the memories, and mapping the judgment onto a response represented by one of the options provided. Below, a discussion of potential errors in college student self-reported gains is organized into these categories.

Comprehension

The language used in self-reported gain items, such as "thinking critically and analytically," is sometimes quite vague (Bowman, 2010a; Porter, 2011). Do students know what this phrase means? If so, do they all share the same definition? And are these definition(s) the same as the researchers' definition(s)? Even experts disagree considerably on the meaning of commonplace terms such as "intelligence" (e.g., Sternberg & Detterman, 1986), so it is reasonable to assume that students may also have different interpretations of terms used in self-reported gain items, such as "critical thinking skills," "general knowledge," and "leadership abilities" (Higher Education Research Institute [HERI], 2011, p. 1). This concern is further complicated by the fact that substantial cross-cultural differences exist on what constitutes complex thinking, interpersonal relationships, and even how a person defines oneself (for reviews, see Kitayama & Cohen, 2009; Markus & Kitayama, 1991; Nisbett, 2003). Thus, students from divergent cultural backgrounds may have systematically different interpretations of a given item. Moreover, some items are double-barreled in that they ask about two concepts at once. For example, if students are asked to report gains in "being an informed and active citizen" (National Survey of Student Engagement [NSSE], 2013, p. 6), then they might have a difficult time knowing how to respond, especially if they have become much more informed but not necessarily more active.

Retrieval and Judgment

The cognitive demands required to provide accurate self-reported gains are substantial. Ideally, students would estimate their own current skills or attributes, estimate their previous skills or attributes, and then have some means for directly comparing the two. However,

students generally do not follow this process; instead, they estimate their current skills and attributes and then attempt to determine whether or how these have changed over time (Ross, 1989). This distortion of the ideal process can lead to substantial errors, because students' estimates are biased toward their lay theories of change and stability over the lifespan. As Ross explains, most people think that their skills generally increase over time (with the exception of very late in life), whereas they think that their attitudes are quite stable. As a result, consistent with these lay theories, people tend to overestimate how much their skills and abilities have changed, yet underestimate how much their attitudes have changed (Conway & Ross, 1984; Goethals & Reckman, 1973; Markus, 1986; McFarland & Ross, 1987).

Interestingly, students may be reasonably accurate when estimating their current skills. Some early research found high correlations between self-reported knowledge and objectively tested knowledge (Berdie, 1971; Pohlmann & Beggs, 1974), and other studies found that self-reported and objectively tested skills on the same academic subject load onto the same factor within structural equation models (Pike, 1995, 1996). Moreover, a recent meta-analysis found a moderate relationship between objective measures of one's *current* knowledge level and self-assessments of knowledge (r = .34), whereas there was no relationship when examining *increases* in self-perceived and actual knowledge (r = .00; Sitzmann, Ely, Brown, & Bauer, 2010). Thus, the errors on self-reported gains may primarily occur not because of students' inadequate self-knowledge of their current attributes, but because they cannot or do not use adequate processes to estimate their growth over time.

Two additional biases may be considered to involve both difficulties with retrieval and failures to judge the adequacy of one's memories. Halo error occurs when students' perceptions of overall growth and development unduly influence their judgment of growth in specific domains. In a classic experimental example, Nisbett and Wilson (1977b) found that students were quite fond of a professor's European accent when he acted warm and friendly in a videotaped interview, whereas other students were annoyed by the same professor's accent when they saw him acting cold and distant in a different interview. Pike (1993) also observed direct evidence of halo error in self-reported gains when seniors reported on their overall collegiate experience. Other studies have provided indirect evidence by finding low correlations among longitudinal gains on various constructs, but moderate to high correlations among self-reported gains, which suggests that the interrelationships among self-reported gains may be inflated (Bowman, 2010b; Bowman & Brandenberger, 2010). Pike (1999) further demonstrated that halo error may account for up to 75% of the explained variance in self-reported gains among first-year students.

It is reasonable to assume that students may also have different interpretations of terms used in self-reported gain items, such as "critical thinking skills," "general knowledge," and "leadership abilities."

In addition, Pascarella (2001) argued that students may differ in the extent to which they perceive their educational experiences as beneficial; these chronic dispositions toward reporting (or not reporting) growth may also constitute an important source of error. He suggests that controlling for students' perceived gains during high school will largely or entirely correct for this error in college self-reported gains, but this practice has rarely been employed in higher education research. Recent studies have found that high school self-reported gains are at least moderately correlated with college self-reported gains (Bowman & Hill, 2011; Seifert & Asel, 2011) and that the results of regression analyses sometimes depend upon whether high school gains are included as a control variable (Seifert & Asel, 2011).

Response

Biases may also occur when students are asked to select a response option. On the NSSE, when reporting how much students' "experience at this institution contributed to [their] knowledge, skills, and personal development," the response options are "very much," "quite a bit," "some," and "very little" (2013, p. 6). All four of these categories are at least implicitly positive—and they are treated as positive in statistical analyses—so students are unable to state that they have not changed at all or that they declined. On the Cooperative Institutional Research Program (CIRP) College Senior Survey, students' response options for changes in their knowledge, skills, and understanding were "much stronger," "stronger," "no change," "weaker," and "much weaker" (HERI, 2011, p. 1). The CIRP scale eliminates some of the problems apparent on the NSSE scale, but only two options are available for reporting positive growth, which could lead to range restriction. Perhaps more importantly, the categories

on both surveys are quite vague. Do students draw similar distinctions between "quite a bit" and "very much" or between "stronger" and "much stronger"? The results from an older study on students' perceptions of college experience frequency descriptors (Pace & Friedlander, 1982) may be informative. When asked about the frequency of making appointments to see faculty members, 21% of students thought that the term "very often" meant more than once a week, whereas 33% of students thought that this meant 1-2 times a month, and a small percentage of students (2%) thought that this meant 1-2 times per year. Clearly, students can assign very different meanings to such descriptors.

Moreover, students may select a response category that portrays them in an overly positive light. For instance, a socially desirable response would be to say that they have gained a great deal while in college; the unappealing alternatives are to say that they have gained very little, not at all, or even regressed. Indeed, social desirability scales are significantly associated with college student self-reported gains (Bowman & Hill, 2011; Gonyea & Miller, 2011), and this relationship persists even when controlling for self-esteem, college satisfaction, and other potential confounding variables (Bowman & Hill, 2011).

Additional Problems and Processes

People tend to overestimate how much their skills and abilities have changed, yet underestimate how much their attitudes have changed.

As Krosnick (1991) explains, survey respondents are likely to become increasingly fatigued, disinterested, and distracted as they continue to take a survey. As a result, participants expend less energy (if any) on each of Tourangeau et al.'s (2000) four steps; Krosnick refers to this suboptimal responding as "satisficing." Self-reported gains may induce satisficing—particularly if they are included later in the survey—because these items require a great deal of cognitive effort, involve responses for which students likely do not have a preconceived answer, and often appear in succession with other such items that use the same response scale. Indeed, Barge and Gehlbach (2012) showed that satisficing is quite common when reporting college self-reported gains and that this tendency may substantially and adversely affect survey results (also see Chen, 2011).

Going a step further, Porter (2013) argues that a belief-sampling model of survey response more adequately captures students' thinking when considering their own growth. That is, instead of recalling actual memories and frequencies of events, students retrieve a host of beliefs, feelings, impressions, values and judgments (collectively referred to as "considerations") that are relevant to the question. The specific set of considerations that students retrieve is somewhat arbitrary and is based on what is accessible in that particular time and context. Porter offers an example of what this process might look like:

> Consider a student in a quantitatively-oriented major who is asked how her college experiences have contributed to her development in analyzing quantitative problems. Multiple considerations then enter her mind: memories of lectures from a statistics class; memories of having possibly worked on problem sets with other groups of students; a general impression that she [is] adept at math, based in part on her experiences in high school. These multiple, positive considerations then lead her to conclude that she has gained considerably in analyzing quantitative problems while in college. It is important to note that these considerations could easily be generated by a student, *but that none of them have anything to do with how much a student has learned while in college*. Because considerations that come into mind are a "haphazard assortment," it is clear that many, if not all, of the considerations that enter a student's mind will be related to their educational experiences, but not necessarily to how much they have actually learned in a specific content area. (p. 210, emphasis in original)

Of course, this hypothetical student may be "correct" in the self-assessment of her changes in quantitative skills, but the widespread use of this approach will be largely problematic for drawing conclusions about student growth in the aggregate. Porter tested several hypotheses regarding students' mental processes when reporting their own gains, and the results were quite consistent with predictions from the belief-sampling model. In addition, Bowman and Schuldt (in press) found that students' self-reported gains were higher when these appeared toward the beginning of a questionnaire than when presented toward the end (after reporting their college experiences), which also suggests that the mental availability of certain events likely influences student responses.

Conditions Associated with the Validity of Self-Reported Gains

The preceding discussion paints a rather gloomy picture of the use of self-reported gains as indicators of student learning and growth. However, there is reason to believe that this picture may be somewhat more optimistic under certain conditions. The validity of self-reported gains is substantially determined by the extent to which the outcome is salient and accessible to students. In their classic review, Nisbett and Wilson (1977a) argued that people generally have minimal access to their higher-order cognitive processes, and people's "introspection" on these processes is generally based on their lay theories of cognition. Psychologists have made similar arguments more recently about self-knowledge regarding one's own motivations (Wilson, 2002) and even which activities will lead to one's own happiness (Gilbert, 2007). While many people may have difficulty accessing introspective knowledge accurately, some students may be more attuned to their growth (or lack thereof) on a given outcome. For instance, many first-generation university students face considerable difficulties in their academics and social engagement (e.g., Pascarella, Pierson, Wolniak, & Terenzini, 2004; Zwerling & London, 1992), so they may be more aware of their cognitive and interpersonal growth. Consistent with this view, the correspondence between self-reported and longitudinal gains is greater among first-generation students than among other students (Bowman, 2010a, 2011b).

Instead of recalling actual memories and frequencies of events, students retrieve a host of beliefs, feelings, impressions, values and judgments (collectively referred to as "considerations") that are relevant to the question.

Moreover, students may be much better at estimating their growth on some outcomes than on others. For example, foreign language skills are largely developed through salient formal and informal experiences, and students receive regular feedback on these skills through course grades, instructor comments, and their (in)ability to communicate effectively. In contrast, leadership skills are harder to define, less subject to concrete feedback, and are not often quantified in terms of objective performance. A recent meta-analysis suggests that these outcome attributes are important; specifically, the correspondence between cognitive learning and self-assessments of knowledge is greater when participants are provided external feedback and when they have to opportunity to practice making their own self-assessments (Sitzmann et al., 2010). Perhaps for these reasons, the correlations between longitudinal and self-reported gains are virtually zero for abstract cognitive skills (which generally are not subject to direct feedback or frequent self-assessment), whereas these correlations are somewhat higher for non-cognitive attributes, such as attitudes, interpersonal skills, and intrapersonal knowledge (Bowman, 2010b, 2011b; Sitzmann et al., 2010).

Similarly, the phrasing of self-reported gain items may also affect their validity. For instance, even if students actually knew how much their cognitive skills had changed over time, it is unlikely that all students would have the same interpretation of "thinking critically and analytically," because this construct is quite broad and it contains academic jargon (Porter, 2011). Moreover, students' interpretations of the meaning of some outcomes might differ systematically. For example, "leadership skills" may connote something very different for White, middle-class North Americans (whose cultural contexts generally value individualism and uniqueness) than for Asians and Asian Americans (whose cultural contexts generally value collectivism and consensus; see Nisbett, 2003; Triandis, 1989). These problems can be remedied, in part, by using concrete language that has a similar meaning across diverse groups of students.

For example, foreign language skills are largely developed through salient formal and informal experiences, and students receive regular feedback on these skills through course grades, instructor comments, and their (in)ability to communicate effectively. In contrast, leadership skills are harder to define, less subject to concrete feedback, and are not often quantified in terms of objective performance.

The validity of self-reported gains also depends, in part, upon students' year in college. Several studies have indicated that biases in self-reported gains (e.g., socially desirable responding) appear to be greater among first-year undergraduates than among advanced undergraduates (Bowman & Hill, 2011; Pike, 1999; Seifert & Asel, 2011). This pattern may occur for multiple reasons. First, developmental research suggests that self-perceptions generally become more accurate among older children (Harter, 1999), and similar developmental processes may be driving these differences among traditional-age college students. Second, when students are in their last term of their undergraduate education, they may reflect upon their university experiences and how they have changed while attending college. As a result, these students may provide more accurate responses because they have previously considered their growth over time as opposed to providing answers that simply seem plausible (see Krosnick, 1991).

Although self-reported gains are more trustworthy under certain circumstances than in others, longitudinal studies are certainly preferable to cross-sectional studies for drawing inferences about change over time.

The final three attributes relate to issues that were discussed previously. First, self-reported gains will be more valid when an appropriate response scale is used; allowing students to say that a desired attribute did not change or has diminished is generally preferable. As one illustration, when graduating students were asked to provide self-reported gains in their religious beliefs and convictions (and were provided this full set of response options), almost half reported no change during university, and about 14% reported decreases (Lee, 2002). Second, social desirability also plays a role in the accuracy of self-reported gains. The prevalence of socially desirable responding may depend upon the phrasing of the instructions and items as well as the nature of the outcome itself. For example, it is probably less "threatening" for college students to report that they have not become more religious (which is not central to the mission and intended outcomes of most colleges and universities) than to report that their problem-solving skills have not changed. Third, halo error can be more problematic in certain circumstances. Some outcomes appear to be more susceptible to halo error than others; Pike (1993) found that self-reported gains in "personal development" (e.g., intrapersonal skills, self-directed learning) were much more strongly influenced by halo error than self-reported gains in quantitative skills and in understanding arts and cultures. The latter outcomes are fairly specific and not directly related to many students' undergraduate experiences, which likely explains why they are less conflated with general perceptions of growth.

Implications for Assessment in Student Affairs

The following suggestions are provided specifically with student affairs practitioners in mind, but these recommendations may also be useful for institutional researchers, higher education researchers, and others who want to design effective college student assessments.

1. Use longitudinal methods whenever possible. Although self-reported gains are more trustworthy under certain circumstances than in others, longitudinal studies are certainly preferable to cross-sectional studies for drawing inferences about change over time. After asking about self-reported gains for the past 20 years, the Cooperative Institutional Research Program (CIRP) removed these items from its 2013 Your First College Year and College Senior Surveys (see HERI, 2013), which suggests that this organization may have doubts about the usefulness of these items. Because the responses to these CIRP surveys are paired with The Freshman Survey—and all three surveys ask participants to report their current levels of various skills and attributes—CIRP datasets can still assess longitudinal changes during college.

2. Use specific language and multiple items to measure each student outcome. This recommendation actually combines two suggestions, but these are sufficiently related that they should be discussed together. For instance, asking students directly about "leadership skills" provides problems regarding both the ambiguity of language and the multidimensionality of this complex construct; in short, what exactly is meant by "leadership"? This problem can be remedied by providing items that measure behaviors, attitudes, values, and tendencies that exemplify various aspects of leadership. The original Socially Responsible Leadership Scale (SRLS) contained 104 items that indicate eight leadership constructs (Tyree, 1998). While this instrument constitutes an extreme example of the number of items (and subsequent versions of the SRLS contain fewer items), it illustrates the extent to which a complex concept can be measured in detail when it is the primary focus of a research or assessment project.

3. Never ask students to self-report their cognitive growth. There still may be some hope that a well-designed questionnaire can yield accurate estimates of student gains on some affective outcomes. However, self-reported and longitudinal assessments of cognitive outcomes provide such strongly divergent findings that these self-reports appear completely untrustworthy. As described earlier, standardized examinations and authentic assessments (e.g., portfolios or rubrics of student work) are likely the most effective means for assessing cognitive and academic growth.

4. Give pretests and posttests for content-based workshops and programs. As a way of exploring learning outcomes within a program or workshop, students could take a closed- or open-ended quiz on key concepts. This approach could be successful for a longer program (e.g., professional development over a semester), and a short quiz could also be useful for one- or two-hour workshops (e.g., regarding career planning). For the short version, some people may be skeptical of using a single quiz for both the pretest and posttest, because students'

responses may exhibit practice effects or students may be overly attentive to these specific pieces of information. If this seems problematic, two versions of the test could be created; half of the students complete Version A in the pretest and Version B in the posttest, and the other half of students would complete Version B and then Version A.

5. Collaborate across campus to conduct large-scale assessments. Coordinating efforts across departments, units, and divisions (including student affairs and academic affairs) can result in comprehensive assessments that would not otherwise be possible. For instance, students who take a critical thinking examination and/or other in-depth instruments might also report their involvement in various curricular and cocurricular activities so that one can determine whether these experiences predict performance and growth. This approach may also have the benefit of reducing survey fatigue, which has helped contribute to dramatic recent reductions in survey response rates (Adams & Umbach, 2012; Pew Research Center, 2012).

This emerging research on self-reported gains has strengthened my belief that having poor quality data is highly problematic and potentially misleading. The predictors of college student self-reported gains and longitudinal growth on the same construct differ considerably, and this divergence is sometimes systematic and even predictable.

Conclusion

A few years ago, a colleague and I had several discussions about whether it is preferable to have poor quality data or no data at all. This emerging research on self-reported gains has strengthened my belief that having poor quality data is highly problematic and potentially misleading. The predictors of college student self-reported gains and longitudinal growth on the same construct differ considerably (Bowman, 2010b, 2012; Bowman & Brandenberger, 2010), and this divergence is sometimes systematic and even predictable (Bowman, 2011a; Conway & Ross, 1984; Porter, 2013). Therefore, higher education practitioners and administrators can make faulty decisions about programs and practices if they rely too strongly upon students' subjective perceptions of learning and growth. Student affairs professionals face a host of circumstances that make them more likely to use this type of outcome assessment, so they must be particularly diligent about avoiding the problems associated with perceived growth. Although it is certainly more challenging and expensive to collect high-quality, longitudinal data on student outcomes, the long-term benefits will generally outweigh the costs.

AUTHOR'S NOTE:

I thank Vivienne Felix for her feedback on an earlier version of this manuscript.

References

Adams, M. J. D., & Umbach, P. D. (2012). Nonresponse and online student evaluations of teaching: Understanding the influence of salience, fatigue, and academic environments. *Research in Higher Education, 53*, 576-591.

Anaya, G. (1999). College impact on student learning: Comparing the use of self-reported gains, standardized test scores, and college grades. *Research in Higher Education, 40*, 499-526.

Astin, A. W., & antonio, a. l. (2012). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education* (2nd ed.). New York, NY: Rowman & Littlefield/American Council on Education.

Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice.* San Francisco, CA: Jossey-Bass.

Barge, S., & Gehlbach, H. (2012). Using the theory of satisficing to evaluate the quality of survey data. *Research in Higher Education, 53*, 182-200.

Berdie, R. F. (1971). Self-claimed and tested knowledge. *Educational and Psychological Measurement, 31,* 629-636.

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). San Francisco, CA: Jossey-Bass.

Bowman, N. A. (2010b). Can 1st-year college students accurately report their learning and development? *American Educational Research Journal, 47*, 466-496.

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). San Francisco, CA: Jossey-Bass.

Bowman, N. A. (2011b). Validity of self-reported gains at diverse institutions. *Educational Researcher, 40*(1), 22-24.

Bowman, N. A. (2012, March). *Measuring student growth? The meaning and interpretation of self-reported gains*. Paper presented at the annual meeting of the American College Personnel Association, Louisville, KY.

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). Charlotte, NC: 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). San Francisco, CA: Jossey-Bass.

Bowman, N. A., & Schuldt, J. P. (in press). Effects of item order and response options on college student surveys. In N. A. Bowman & S. Herzog (Eds.), *Methodological advances and issues in studying college impact* (New Directions for Institutional Research). San Francisco, CA: 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*, 270-290.

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*, 659-674.

Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47*, 738-748.

Gilbert, D. (2007). *Stumbling on happiness*. New York, NY: Vintage.

Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology, 9*, 491-501.

Gonyea, R. M., & Miller, A. (2011). Clearing the AIR about the use of self-reported gains in institutional research. In S. Herzog & N. A. Bowman (Eds.), *Validity and limitations of college student self-report data* (New Directions for Institutional Research, no. 150, pp. 99-111). San Francisco, CA: Jossey-Bass.

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

Harter, S. (1999). *The construction of the self: A developmental perspective.* New York, NY: Guilford.

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

Higher Education Research Institute. (2011). *2011-2012 College Senior Survey*. Retrieved from http://www.heri.ucla.edu/researchers/instruments/FUS_CSS/2012CSS.PDF

Higher Education Research Institute. (2013). *Survey instruments, codebooks, & participation history*. Retrieved from http://www.heri.ucla.edu/researchersToolsCodebooks.php

Kitayama, S., & Cohen, D. (2009). *Handbook of cultural psychology*. New York, NY: Guilford.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213-236.

Kuh, G., & Ikenberry, S. (2009). *More than you think, less than we need: Learning outcomes assessment in American higher education*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Lee, J. J. (2002). Religion and college attendance: Change among students. *Review of Higher Education, 25*, 369-384.

Markus, G. B. (1986). Stability and change in political attitudes: Observed, recalled and explained. *Political Behavior, 8*, 21-44.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.

McFarland, C., & Ross, M. (1987). The relation between current impressions and memories of self and dating partners. *Personality and Social Psychology Bulletin, 13*, 228-238.

National Survey of Student Engagement. (2013). *National Survey of Student Engagement 2013* (U.S. English Version). Retrieved from http://nsse.iub.edu/pdf/survey_instruments/2013/NSSE%202013%20US%20English.pdf

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently…and why.* New York, NY: Free Press.

Nisbett, R. E., & Wilson, T. D. (1977a). Telling more than we know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

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

Pace, C., & Friedlander, J. (1982). The meaning of response categories: How often is occasionally, often, and very often? *Research in Higher Education, 17*, 267-281.

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

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *Journal of Higher Education, 75*, 249-284.

Pew Research Center. (2012). *Assessing the representativeness of public opinion surveys.* Washington, DC: Author.

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

Pike, G. R. (1995). The relationship between self reports of college experiences and achievement test scores. *Research in Higher Education, 36*, 1-21.

Pike, G. R. (1996). Limitations of using students' self-reports of academic development as proxies for traditional achievement measures. *Research in Higher Education, 37*, 89-114.

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

Pohlmann, J., & Beggs, D. (1974). A study of the validity of self-reported measures of academic growth. *Journal of Educational Measurement, 11*, 115-119.

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

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

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

Seifert, T. A., & Asel, A. M. (2011). The tie that binds: The role of self-reported high school gains in self-reported college 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. 59-72). San Francisco, CA: Jossey-Bass.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*, 169-191.

Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco, CA: Jossey-Bass.

Sternberg, R. J., & Detterman, D. K. (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.

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

Triandis, H. C. (1989). The self and social behavior in differing social contexts. *Psychological Review, 96*, 506-520.

Tyree, T. (1998). *Designing an instrument to measure socially responsible leadership using the social change model of leadership development*. Unpublished doctoral dissertation, University of Maryland-College Park.

Walvoord, B. E. (2010). *Assessment clear and simple* (2nd ed.). San Francisco, CA: Jossey-Bass.

Wilson, T. D. (2002). *Strangers to ourselves*. Cambridge, MA: The Belknap Press of Harvard University Press.

Zwerling, L. S., & London, H. B. (Eds.) (1992). *First-generation students: Confronting the cultural issues* (New Directions for Community Colleges, no. 80). San Francisco, CA: Jossey-Bass.