Benchmarks for scholarly investigations of teaching and learning

Georjeanna Wilson-Doenges and Regan A R. Gurung

University of Wisconsin-Green Bay, Green Bay, Wisconsin, USA

Abstract

The scholarship of teaching and learning (SoTL) is a growing and vibrant field, and with that growth has come some serious reflection on the models we use to conduct SoTL and the robustness of these models. Unfortunately, there is no clear-cut set of benchmarks for robust research. In this article, we identify a continuum of SoTL and demarcate aspirational benchmarks. Whereas qualitative and quantitative data and methods all have a place in SoTL, psychological science provides gold standards of design and analysis. We advocate for similar SoTL standards that are theory based and intentionally designed using the best models for methodological and statistical rigor.

Key words: assessment, learning, SoTL, teaching, university

The scholarship of teaching and learning (SoTL) incorporates a wide array of practices whereby instructors closely examine student learning, systematically and intentionally change pedagogy, assess the results of change, and share their findings. SoTL is now an established area of scholarship, but there are a number of issues within the field (Boshier, 2009; Dewar, Dailey-Hebert, & Moore, 2010). Perhaps the most critical one concerns, 'the legitimacy of the literatures and methods that shape teachers' questions about learning and the kinds of evidence they seek in order to answer them' (Hutchings, Huber, & Ciccone, 2011, p. 8). SoTL practitioners are not always aware of (or do not always use) adequately rigorous research designs or analyses (Dunn, McCarthy, Baker, Halonen, & Boyer, 2011). Different fields have competing definitions of excellence regarding scholarship (Lamont, 2009), and SoTL does not have a clear-cut definition of excellence at all. In this article, we rectify this shortcoming and outline SoTL benchmarks of rigor to shape, guide, and motivate this burgeoning field.

Clarifying terminology

Before we define rigorous SoTL we need to first be clear on what it is. Potter and Kustra (2011, p. 2) define the SoTL as:

the systematic study of teaching and learning, using established or validated criteria of scholarship, to understand how teaching (beliefs, behaviours, attitudes, and values) can maximize learning, and/or develop a more accurate understanding of learning, resulting in products that are publicly shared for critique and use by an appropriate community.

Boyer (1990) popularised the term 'scholarship of teaching'. Huber and Hutchings (2005) further developed SoTL and describe it as 'viewing the work of the classroom as a site for inquiry, asking and answering questions about student learning in ways that can improve one's own classroom and also advance the larger profession of teaching' (p. 1). SoTL also has a well-established history in the discipline of psychology (Gurung, Ansburg, Alexander, Lawrence, & Johnson, 2008; Gurung & Schwartz, 2013; Tomcho et al., 2008; Worrell et al., 2010).

Definitions and distinctions become important when the degree of assessment varies. The most important case is in the difference between SoTL and scholarly teaching (Potter & Kustra, 2011; Smith, 2012). Scholarly teaching involves reflecting on a classroom issue, deciding to implement a course change, consulting the relevant literature from the knowledge base of teaching and learning in higher education, gathering baseline information, making a change, and measuring the result of this change. Scholarly teachers create course design portfolios to document their systematically collected observations for further reflection and course modification (Richlin, 2006). The portfolios allow for the identification of problem areas, and in contrast to the

instructor who may also informally analyse student problem areas, the scholarly teacher takes pains to venture into the pedagogical literature in search of solutions. This use of an organised methodology and literature review is primarily conducted for the instructors' own benefits, to make their next class better. The results of their reflections and course modifications may be used for merit or promotion but are not submitted for publication or shared with peers. Conducting formal investigations of teaching and learning (regardless of method), placing the results in the context of relevant published pedagogical literature, and then submitting it for peer review and subsequent publication catapult the scholarly teacher to a contributor to the SoTL literature.

SoTL goes beyond just assessment of learning (what all instructors should be doing) as it involves an instructor making an intentional change in pedagogy and then assessing *the effectiveness of that change* on student learning. The change could be in how material is presented, forms of active learning used in class, the extent to which technology is used or not used, the type of assignments used, or alterations in classroom policy. Making no changes but just establishing measures of learning and examining grade distributions is basic assessment and par for the course. Both scholarly teaching and SoTL involve much more effort on the instructor's part.

Major types of SoTL

SoTL is now practised around the world (Gurung & Landrum, 2012b), but there is no general agreement on the assessment approach or the specific methodologies that constitute SoTL. There are three major types of SoTL work (McKinney, 2007; Nelson, 2003; Weimer, 2006):

1. reports on particular classes or courses: gather qualitative or quantitative data from students before and after some change in teaching practice
2. comparisons of courses and students: quantitative and/or qualitative comparisons of multiple courses or sections with a shared measurement instrument or similar instruments
3. syntheses of extant SoTL: quantitative meta-analyses of similar studies (e.g., Hattie, 2009)

Within these major categories of work, there are also myriad ways of measuring classroom learning and teaching (Angelo & Cross, 1993; Maki, 2011). Although psychologists tend to prefer quantitative methods, a number of qualitative methods enable observers to assess concurrently broader dimensions of learning that lend themselves to narrative interpretation. For example, a large number of faculty use 'think alouds' where learners orally report on their thought processes as they tackle an assignment or study (Hutchings

et al., 2011). Relatedly, using the lesson study methodology, groups of faculty work together on the design of a 'lesson', a specific class period or lecture, and then observe student reactions to the instruction in person and using videotape and qualitative and quantitative measures (Cerbin, 2011).

In general, instructors can use experimental and quasi-experimental designs to gather both quantitative and rich qualitative data. There are many ways of researching whether students' learning improves as a result of changes instructors make to their teaching. Robust SoTL challenges teachers to get into their students' heads and as such requires both *more* evidence of learning than traditional assessment, and *different kinds* of evidence (Hutchings et al., 2011). As any research methods textbook intones, some designs and analyses are better than others. We use the spectrum of research designs to lay the foundation for benchmarks of excellence for SoTL.

SETTING SoTL STANDARDS

SoTL should entail methodologically rigorous scholarly work conducted to enhance teaching and advance learning. In addition to clear definitions of SoTL (Gurung & Schwartz, 2010; Potter & Kustra, 2011), there are general criteria of scholarship that provide us with a place to start. For example, Glassick, Huber, and Maeroff (1997) provide criteria to serve as a basis for rigor in SoTL. In general, SoTL should:

- have clear objectives and purposes
- include appropriate engagement with relevant, current literature in the field
- utilise appropriate methods that relate to the stated goals
- show a significant or likely impact on learning
- show effective communication requiring the dissemination of ideas and findings to the wider academic community
- entail reflective critique where 'the scholar thinks about his or her work, seeks the opinion of others, and develops his or her learning over time' (Glassick et al., 1997, p. 5)

Although a good start, psychological science can take things a step further and refer to the 'gold standard' of our field to be sure that there is a 'greater likelihood that trustworthy information will be created and disseminated' (Potter & Kustra, 2011, p. 2). A special issue on SoTL in the *Teaching of Psychology* sets the stage for establishing the gold standard in our field suggesting experimental designs with random assignment where possible should be the foundation of quality SoTL research in psychology (LoSchiavo, Shatz, & Poling, 2008). We advance the recommendations of the SoTL special issue and suggest a basic continuum for

establishing SoTL rigor described later. Each level features different design or analysis characteristics. Level 1 represents a good entry into SoTL but with methodology that is weaker than higher levels. Level 2 shows solid SoTL work but still has some methodological shortcomings. Level 3 represents the gold standard and uses the best and most rigorous methodology to collect quality SoTL data. We acknowledge that achieving the gold standard in SoTL has practical and ethical issues and may be difficult when working within university constraints. Nonetheless, we believe that the continuum can still provide aspirational guidelines for SoTL research.

SoTL BENCHMARKS: LEVEL 1

Theory is generated after the data are collected and analysed

Data are collected, the instructor realises that there are some interesting results, and then goes to the literature to find theoretical support rather than basing the original data collection on theory found prior to data collection.

No comparisons

Cross-sectional research with no pretest or control group (i.e., posttest only, no control) does not enable comparisons. Descriptive data are collected on the dependent variable at one point in time with no ability to make comparisons with a baseline score for the class or to another class that has not been exposed to the SoTL independent variable.

Non-equivalent comparison groups

Comparisons are being made on the dependent variable between the class exposed to the SoTL independent variable and another class, but the comparison class is not equivalent (i.e., vastly different in size, demographics, instructor, or content). Non-equivalence reduces the validity of such comparisons.

Sample sizes less than 25 or too small to have adequate power

When analysing data using inferential statistics, sample size, in part, determines the reliability and the power of the tests. Using small sample sizes often does not meet the basic assumptions of typical inferential statistics such as analysis of variance (ANOVA) or regression. In addition, power (the ability to detect significance) is often compromised by too small of a sample size creating a higher probability of type II error.

Self-report of learning is the only dependent variable

Self-report measures are a part of our understanding of student perceptions of learning, but alone, they are inad-

equate to assess actual learning (Tomcho & Foels, 2008). Measures such as test or quiz scores, grade point average (GPA), and other learning assessments should be included as dependent variables rather than self-report questions alone (i.e., 'Do you think you learned more using flash cards to study for the test?').

Retrospective reports of attitudes/opinions

Asking students to recall opinions or attitudes retrospectively is not as accurate as assessing at the time of learning. Asking students at the end of the semester to provide an opinion about how much they thought they learned from multiple teaching strategies over the course of the semester can be affected by such memory problems as the recency effect or forgetting.

Multiple bivariate hypothesis tests with no adjustment of the level of significance (α)

When conducting multiple comparisons of dependent variables between groups or within groups using bivariate models, an adjustment to the level of significance must be made to accurately represent the probability of type I error. For instance, if running five independent-sample t-tests, using $\alpha = 0.05$ for each test inflates the probability of finding significance when it is not actually there. To correct this problem, a Bonferroni adjustment could be made to the family-wise error rate.

Qualitative data that are not of adequate depth

Using qualitative methodology without adequate depth reduces the validity of the information gathered. For instance, if using narrative interview to understand students' conceptions of difficult concepts in a statistics class, asking questions that do not allow for students to explore their personal feelings of fear or anxiety about math in general may limit the full understanding of the students' thoughts and feelings about those statistical concepts.

SoTL BENCHMARKS: LEVEL 2

Theory and literature known before data are collected

Understanding and incorporating prior research and theory into the conception of a SoTL project can improve the quality of the results. For instance, finding and using an already established, reliable, and valid scale to measure a dependent variable increases the quality of that measurement.

Pretest–posttest experimental design using appropriate analyses (i.e., repeated measures)

Conducting a pretest provides a comparison of your dependent variable at baseline. Using a within-groups statistical test

such as repeated-measures ANOVA can provide information about change in learning due to the SoTL intervention.

Experimental design with equivalent control group (i.e., at least matched)

Another approach to making comparisons is to have a control group who was not exposed to the SoTL independent variable. In most SoTL research, accomplishing random assignment is not feasible or ethical, but having a control group that is as equivalent as possible increases the validity of the measurement of change due to the SoTL intervention. For example, Elicker, O'Malley, and Williams (2008) used two different instructors teaching the same course where one instructor used a standard website, and the other used an enhanced WebCT website. They chose the two classes by selecting two sections of the same course but with two different instructors. They chose instructors whose course evaluations were not significantly different instead of having the same instructor teach both sections of the course because they were concerned about the burden on instructors and the possible biasing of students reactions. Students in these two classes were not significantly different (equivalent) in terms of their ACT scores, GPA, age, and number of days absent.

Established reliability and validity

Using already reliable and validated measures for your dependent variables improves the quality of your measurement. Establishing and reporting reliability and validity of your measurement are important in evaluating the quality of the results.

Acceptable sample size for analysis (at least 25 and at least 10 per level of the independent variable—Sometimes more if needed)

Minimum sample size requirements to meet the assumptions of most inferential statistics include sample sizes of 25 or greater and, in multivariate analyses, having at least 10 participants for each measured independent variable. If these assumptions are violated, the results from these analyses are less reliable.

Simple multivariate analyses (i.e., control for self-reported GPA, year in school, etc.)

Rather than conducting multiple bivariate analyses between one independent and one dependent variable at a time, combining variables into a simple multivariate model increases the reliability of the results. Multivariate models can decrease the probability of making a type I error and increase the validity of the results by controlling for contextual factors.

Qualitative data that are of adequate depth

Using qualitative methodology with adequate depth increases the validity of the information gathered. Understanding learning from the students' perspective in full increases the understanding of the concepts from those experiencing them. Allowing students to share the full domain of their understanding can expand our conception of how students learn from SoTL interventions.

SoTL BENCHMARKS: LEVEL 3: THE GOLD STANDARD

Theory-based scholarly teaching where the SoTL independent variable and dependent variables are based on a strong literature review

A thorough understanding of the literature can vastly improve the design of a SoTL project by improving on the weaknesses and by capitalising on the strengths of previously published studies. Again, using a criterion (an already established measure) from the literature to measure your dependent variable can also boost the reliability and validity of your study.

Longitudinal design tracking students over time with good response rates and analyses that track change scores

Multiple measures of the dependent variable over time in the semester can help to uncover changes in learning over the semester because of an SoTL intervention. For example, tracking changes in quiz scores after implementing a meta-cognition improvement intervention can track improvements in quiz scores over the whole semester showing that, perhaps, the intervention did not help immediately (e.g., no significant change in quiz score between quiz one and quiz two) but did help in the long term over the semester (e.g., significant improvement in quiz score between quiz one and quiz eight) (for a discussion of metacognition, see Dunn, Saville, Baker, & Marek, in this issue). The repeated measures design works well in course designs that have several similar exams or assignments. In such designs, the key is to identify changes in responses to similar question(s) over time. The measure used repeatedly consists of the same number of questions asked in the same order, and differences in the responses will be taken to indicate changes in knowledge (e.g., understanding diversity; Kernahan & Davis, 2010). Using a statistical technique such as repeated-measures ANOVA can capture these types of within-group longitudinal changes.

True experimental design that includes random assignment

Although difficult to achieve in the classroom setting, striving to collect data using the highest standards in our field

will produce the highest quality SoTL work. Using random assignment will insure the equivalence of comparison groups. For example, in a study testing differences in free-recall versus lecture-only course delivery, Balch (2012) gave all students in an introductory psychology class a pretest then randomly assigned them to two groups. He then asked students in the groups to come to class at different times the following class period. The same course material was presented as either free-recall or lecture depending on the condition to which students were assigned. Students were then given a posttest. Using a 2×2 design with random assignment, Balch was able to show an interaction between instructional condition and time of test illustrating the learning value of the free-recall demonstration.

Double-blind design

Again, although difficult to achieve, having those who deliver and receive the SoTL independent variable be blind to which condition the students are receiving is crucial to removing confounds in the study.

Large sample sizes and established power of the test

Completing a power analysis and determining, then collecting, a large enough sample to meet or exceed the sample size recommendation establish good power of the test.

Samples taken from more than one institution with some diversity

Although we can learn a lot from single classes and institutions, we can learn more by collecting nationwide samples to increase the diversity of participants and viewpoints and to increase external validity. In one such example, Brack, Kesitilwe, and Ware (2010) conducted a random national survey of 300 institutions to determine the current state of the undergraduate health psychology course, uncovering national information about the course and its maturation over the last decades.

Proper data screening and statistical analyses that are advanced and multivariate (i.e., regression and structural equation modelling)

Data screening can ensure that your variables meet the important assumptions of the statistical tests that will be used. Data screening also assists in identifying outliers and in highlighting issues with missing data. For example, normality tests, such as the Kolmogorov–Smirnov statistic, should be completed to assess univariate normality for dependent variables. If significant problems are present, data transformation might be a way to address these violations of statistical assumptions. Once transformed, these variables will lead to more accurate statistical results. In addition,

advanced statistical techniques can improve our ability to track changes and to understand complex multivariate models.

High standard of ethics

Insuring that students receive benefits for participation in SoTL research and are not coerced to participate or harmed is imperative in classroom research just as in all research (Gurung, 2012; Swenson & McCarthy, 2012). Providing incentives must be equal and just for all participants. For example, Funk and Dickson (2011) provided extra credit for correct answers to additional questions of various formats on exams. The authors state that the Institutional Review Board of their institution approved giving extra credit because the incentive was 'voluntary and non-punitive' (Funk & Dickson, 2011, p. 274).

Mixed-method approach—Using both qualitative and quantitative data analysed appropriately

Use high-quality qualitative analysis (e.g., NVivo, QSR International (Americas) Inc., Burlington, MA, USA) to analyse qualitative data and combine that information with quantitative data in a single analysis. For example, in a study of intercultural competency training, Knott et al. (2013) used both quantitative analysis of survey data and qualitative thematic analysis of open-ended responses to find support for an increase in cultural competency through skill development.

ADDITIONAL CONSIDERATIONS

In addition to the aforementioned continuum of research methodology and analyses, there are other key elements that must be kept in mind to conduct meaningful SoTL.

First, to best conduct SoTL, one should have a clear picture of one's learning outcomes. Student learning outcomes (SLOs) are the 'knowledge, skills, attitudes, and habits of mind that students take with them from a learning experience' (Suskie, 2009, p. 117). One can think of SLOs as how SoTL researchers operationally define their dependent variables to measure learning (Gurung & Landrum, 2012b). The systematic measurement of SLOs provides the evidence that learning occurred. Well-written SLOs possess three characteristics: performance outcomes, contextual conditions, and behavioural criteria (Mager, 1962; Marzano, 2009). First, an SLO should describe what the student should know or be able to do after the learning experience in precise terms—analogous to a good operational definition. Second, an SLO may describe the conditions or context in which the demonstration of knowledge and/or skills is to occur—that is, the context. Third, educators should establish a criterion—an

acceptable level or of performance—to help determine if the student has achieved the SLO.

Second, SoTL research should not be restricted to an assessment of a single classroom technique one semester in a small class. Potter and Kustra (2011) focus SoTL research on the 'systematic study' of teaching and learning which is further defined as 'deliberate, planned, intentional, occurring over time and refined as necessary' (Potter & Kustra, 2011, p. 2). There are outlets for basic assessment of teaching techniques, but the SoTL should be held to a higher standard of deliberate, well-planned, programmatic, and designed research that should extend, if possible, beyond a semester and a single class. It is the intentionality that sets the bar for quality. Buskist, Carlson, Christopher, Prieto, and Smith (2008, p. 273) explain, 'Intentionality in action permits a clearer distinction between causal effect and chance effect with respect to student learning.' As shown previously, we should strive to conduct intentional investigations that capture as much of the contextual elements of teaching and learning as possible while using the most methodologically sound yet ethical designs possible.

Third, conducting meaningful SoTL must entail robust assessment (Walvoord, 2004). Although scholarly teachers are assessing the outcomes of the changes they make to their teaching, the results may not go any further (i.e., be published or presented). The assessment may consequently, implicitly or explicitly, not be at the same rigorous level as if the results were to be publically disseminated. Compare a magazine article with a journal article—a different tone and level of credibility is ascribed depending on the outlet. Not having to measure up to the standards required of peer review could affect the extent to which the assessment is undertaken. Peer-reviewing then ensures the quality of the SoTL work. 'The scrutiny of peer review of SoTL will ensure that flaws are minimised, loopholes are identified, and methodologies are watertight' (Gurung & Schwartz, 2009, p. 15).

Finally, there is also a greater need for better, more integrated, theoretical work (Gurung & Schwartz, 2010) to complement and situate the methodological and statistical rigor advocated previously. Hutchings (2007) noted that 'the role of theory in the SoTL is the elephant in the room' (p. 1)'. Reflecting on the 2007 International Society for the Scholarship of Teaching and Learning annual meeting, Hutchins noted how many presentations lacked a theoretical base. We need to work harder to take basic research in relevant areas and to apply it to teaching and learning. For example, cognitive psychologists and social psychologists are nicely taking theoretically driven lab work and are applying it to the classroom (Bjork & Bjork, 2011; Gurung & Burns, 2011). Going beyond this call for theory is the need to situate all the myriad studies of pedagogical research in a common context. We are all trying to understand how students learn best.

Whereas tests of individual class activities and techniques are important, it is now time for us to look at the big picture.

We have some helpful models for guidance. In perhaps one of the largest meta-analyses conducted in pedagogical research, Hattie (2009) analysed over 800 meta-analyses of studies relating to achievement and lists 131 factors that influence learning. Similarly, there is now a growing body of literature integrating different variables into a picture of learning. Most recently, Bernstein et al. (2010) and Chew et al. (2010) provide comprehensive pictures of what is known about the processes surrounding teaching and learning and provide general models that can guide future pedagogical research. Similarly, Shell, Brooks, Trainin, Wilson, Kauffman, and Herr (2010 take concepts from the cognitive, motivation, and neurobiological sciences and use them to set out a unique theory of learning. These are exactly the types of endeavours that more pedagogical researchers need to be aware of and use to position their own research.

In conclusion, we hope that we have underscored the vibrancy of the field of SoTL while highlighting some key methodological approaches. The highest form of SoTL should be theory based, have established power, use reliable and valid measures, use robust methodologies, and utilise advanced and multivariate techniques to analyse properly screened data. Living up to the gold standard in our field will enable SoTL research to garner the same respect that our discipline-specific research and publications garner.

REFERENCES

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco, CA: Jossey-Bass.

Balch, W. R. (2012). A free-recall demonstration versus a lecture-only control: Learning benefits. *Teaching of Psychology, 39*(1), 34–37.

Bernstein, D., Addison, W., Altman, C., Hollister, D., Meera, K., Prieto, L. R., . . . Shore, C. (2010). Toward a scientist-educator model of teaching psychology. In D. Halpern (Ed.), *The NCUEP: A blueprint for the future* (pp. 29–46). Washington, DC: APA.

Bjork, E. L., & Bjork, R. (2011). Making things harder on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustration fundamental contributions to society* (pp. 56–64). New York: Worth.

Boshier, R. (2009). Why is the scholarship of teaching and learning such a hard sell? *Higher Education Research and Development, 28,* 1–15. doi:10.1080/07294360802444321

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. San Francisco, CA: Jossey-Bass.

Brack, A. B., Kesitilwe, K., & Ware, M. E. (2010). Taking the pulse of undergraduate health psychology: A nationwide survey. *Teaching of Psychology, 37*(4), 271–275.

Buskist, W., Carlson, J. F., Christopher, A. N., Prieto, L., & Smith, R. A. (2008). Models and exemplars of scholarship in the teaching of psychology. *Teaching of Psychology, 35,* 267–277. doi:10.1080/00986280802373908

Cerbin, B. (2011). *Lesson study: Using classroom inquiry to improve teaching and learning in higher education.* Arlington, VA: Stylus.

Chew, S. L., Bartlett, R. M., Dobbins, J. E., Hammer, E. Y., Kite, M. E., Loop, T. F., . . . Rose, K. C. (2010). A contextual approach to teaching: Bridging methods, goals and outcomes. In D. Halpern (Ed.), *The NCUEP: A blueprint for the future* (pp. 95–112). Washington, DC: APA.

Dewar, J., Dailey-Hebert, A., & Moore, T. (2010). The attraction, value and future of SoTL: Carnegie affiliates' perspective. *Transformative Dialogues: Teaching and Learning Journal, 4,* 1–15.

Dunn, D. S., McCarthy, M. A., Baker, S. C., Halonen, J. S., & Boyer, S. (2011). Understanding faculty reluctance to assess teaching and learning. In D. Mashek & E. Y. Hammer (Eds.), *Empirical research in teaching and learning: Contributions from social psychology* (pp. 143–159). Malden, MA: Wiley-Blackwell.

Elicker, J. D., O'Malley, A. L., & Williams, C. M. (2008). Does an interactive WebCT site help students learn? *Teaching of Psychology, 35,* 126–131.

Funk, S. C., & Dickson, K. L. (2011). Multiple-choice and short-answer exam performance in a college classroom. *Teaching of Psychology, 38*(4), 273–277.

Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the professoriate.* San Francisco, CA: Jossey-Bass.

Gurung, R. A. R. (2012). Consuming scholarship of teaching and learning: Using evidence-based pedagogy ethically. In R. E. Landrum & M. A. McCarthy (Eds.), *Teaching ethically: Challenges and opportunities* (pp. 67–76). Washington, DC: American Psychological Association.

Gurung, R. A. R., Ansburg, P. I., Alexander, P. A., Lawrence, N. K., & Johnson, D. E. (2008). The state of scholarship of teaching and learning in psychology. *Teaching of Psychology, 35,* 249–261.

Gurung, R. A. R., & Burns, K. (2011). The social psychology of teaching and learning. In E. Y. Hammer & D. Mashek (Eds.), *Social psychology and teaching* (pp. 1–31). Malden, MA: Wiley-Blackwell.

Gurung, R. A. R., & Landrum, R. E. (2012a). Using SoTL to improve learning outcomes. In J. Groccia, M. Alsudairy, & W. Buskist (Eds.), *Handbook of college and university teaching: Global perspectives* (pp. 29–44). Thousand Oaks, CA: Sage.

Gurung, R. A. R., & Landrum, R. E. (2012b). Assessment and the scholarship of teaching and learning. In D. Dunn, S. C. Baker, C. M. Mehrotra, R. E. Landrum, & M. McCarthy (Eds.), *Assessing teaching and learning in psychology: Current and future perspectives* (pp. 159–171). Belmont, CA: Cengage.

Gurung, R. A. R., & Schwartz, B. M. (2009). *Optimizing teaching and learning: Practicing pedagogical research.* Malden, MA: Wiley-Blackwell.

Gurung, R. A. R., & Schwartz, B. M. (2010). Riding the third wave of SoTL. *International Journal for the Scholarship of Teaching and Learning, 4*(2), Retrieved from http://www.academics.georgiasouthern.edu/ijsotl/v4n2/invited_essays/_GurungSchwartz/index.html

Gurung, R. A. R., & Schwartz, B. M. (2013). Contributions from psychology: Heuristics for interdisciplinary advancement of SoTL. In K. McKinney (Ed.), *The scholarship of teaching and learning in and across the disciplines.* (pp. 34–53). Bloomington Indiana: Indiana University Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Huber, M. T., & Hutchings, P. (2005). *The advancement of learning: Building the teaching commons.* San Francisco, CA: Jossey-Bass.

Hutchings, P. (2007). Theory: The elephant in the scholarship of teaching and learning room. International Journal for the Scholarship of Teaching and Learning, 1. Retrieved from http://www.academics.georgiasouthern.edu/ijsotl/v1n1/essays/hutchings/IJ_Hutchings.pdf

Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The scholarship of teaching and learning reconsidered: Institutional impact.* San Francisco, CA: Jossey-Bass.

Kernahan, C., & Davis, T. (2010). What are the long-term effects of learning about racism? *Teaching of Psychology, 37*(1), 41–45. doi:10.1080/00986280903425748

Knott, V. E., Mak, A. S., & Neill, J. T. (2013). Teaching intercultural competencies in introductory psychology via application of the Excellence in Cultural Experiential Learning and Leadership model. *Australian Journal of Psychology, 65*(1), 46–53.

Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment.* Cambridge, MA: Harvard University Press.

LoSchiavo, F. M., Shatz, M. A., & Poling, D. A. (2008). Strengthening the scholarship of teaching and learning via experimentation. *Teaching of Psychology, 35,* 301–304. doi:10.1080/00986280802377164

Mager, R. F. (1962). *Preparing objectives for programmed instruction.* Palo Alto, CA: Feardon.

Maki, P. L. (2011). *Assessing for learning: Building a sustainable commitment across the institution.* Sterling, VA: Stylus.

Marzano, R. J. (2009). *Designing and teaching learning goals and objectives.* Bloomington, IN: Marzano Research Laboratory.

McKinney, K. (2007). *Enhancing learning through the scholarship of teaching and Learning.* Bolton, MA: Anker.

Nelson, C. (2003). Doing it: Examples of several of the different genres of the scholarship of teaching and learning. *Journal on Excellence in College Teaching, 14*(2), 85–94.

Potter, M. K., & Kustra, E. (2011). The relationship between scholarly teaching and SoTL: Models, distinctions, and clarifications. International Journal for the Scholarship of Teaching and Learning, 5. Retrieved from http://www.georgiasouthern.edu/ijsotl

Richlin, L. (2006). *Blueprint for learning: Constructing college courses to facilitate, assess, and document learning.* Sterling, VA: Stylus.

Shell, D. F., Brooks, D. W., Trainin, G., Wilson, K. M., Kauffman, D. F., & Herr, L. M. (2010). *The unified learning model: How motivational, cognitive, and neurobiological sciences inform best teaching practices.* Springer, New York.

Smith, R. A. (2012). Benefits of using SoTL in picking and choosing pedagogy. In B. M. Schwartz & R. A. R. Gurung (Eds.), *Evidence based-teaching for higher education* (pp. 7–22). Washington, DC: American Psychological Association.

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

Swenson, E. V., & McCarthy, M. A. (2012). Ethically conducting the scholarship of teaching and learning research. In R. E. Landrum & M. A. McCarthy (Eds.), *Teaching ethically: Challenges and opportunities* (pp. 21–30). Washington, DC: American Psychological Association.

Tomcho, T. J., & Foels, R. (2008). Assessing effective teaching of psychology: A meta-analytic integration of learning outcomes. *Teaching of Psychology, 35,* 286–296.

Tomcho, T. J., Foels, R., Rice, D., Johnson, J., Moses, T. P., Warner, D. J., . . . Amalfi, T. (2008). Review of ToP teaching strategies: Links to students' scientific inquiry skills development. *Teaching of Psychology, 35*(3), 147–159. doi:10.1080/00986280802201976

Walvoord, B. E. (2004). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. San Francisco, CA: Jossey-Bass.

Weimer, M. (2006). *Enhancing scholarly work on teaching and learning: Professional literature that makes a difference*. San Francisco, CA: Jossey-Bass.

Worrell, F. C., Casad, B. J., Daniel, D. B., McDaniel, M., Messer, W. S., Miller, H. L., . . . Zlokovich, M. S. (2010). Promising principles for translating psychological science into teaching and learning. In D. Halpern (Ed.), *The NCUEP: A blueprint for the future* (pp. 129–144). Washington, DC: APA.