THE STATUS OF REFORM IN STATISTICS EDUCATION: A FOCUS ON THE INTRODUCTORY COURSE

Rossi A. Hassad
Mercy College, New York
Rhassad@mercy.edu

For almost two decades, the statistics education community has been actively focused on reform, particularly regarding introductory statistics. Reform is intended to improve student learning outcomes, and in general, this approach is guided by the constructivist philosophy, and emphasizes the use of active learning and student-centered strategies, so as to foster statistical literacy. Nonetheless, formal assessment data on the effectiveness of reform efforts are lacking. As well, there is no recognized consensus on acceptable measures and indicators, in this regard. This paper presents the results of a qualitative systematic review of literature on statistics education, and concludes that the discipline must urgently heed a call for change toward a more organized and scientific approach to measurement, research, and setting of priorities.

INTRODUCTION

For close to two decades, the statistics education community has been actively and formally engaged in reform, with particular attention to the introductory statistics course at the college level (Tishkovskaya & Lancaster, 2012). It is safe to say that the initial charge for reform was the *Cobb Report* (Cobb, 1992), which was seminal in coalescing support for formalizing the discipline of statistics education. Since then we have witnessed a rapidly expanding reform infrastructure, targeting course content, pedagogy, assessment, and integration of technology, toward fostering statistical and quantitative literacy (Cobb, 2013; Rossman & Garfield, 2011; Aliaga et al., 2005).

In general, the reform approach is guided by the constructivist philosophy, which emphasizes the use of active learning and student-centered strategies, aimed at making the introductory statistics course more practical, engaging, and meaningful to students; rather than mathematical (Hassad, 2013). A tremendous amount of financial and other resources is being spent on reform, nonetheless, formal assessment data on the effectiveness of reform efforts are lacking. As well, there is no recognized consensus on acceptable measures of effectiveness, in this regard. Indeed, the literature contains many reports of effective pedagogy but mostly linked to improvement in student attitude toward statistics (Schau, Millar, & Petocz, 2012), with a dearth of empirical evidence of the impact of reform strategies on academic learning outcomes.

While it has been said that statistics education as a discipline has come of age, the lack of evidence of the effectiveness of reform-based instruction is a major barrier to advancing the discipline. Indeed, there is now some consensus regarding a framework for assessment of student learning outcomes for introductory statistics (Pearl et al., 2012), and psychometric instruments for measuring and characterizing teaching practice, attitudes and beliefs (Hassad, 2011; Zieffler et al., 2012), technology integration (Hassad, 2013), as well as statistical reasoning and conceptual understanding (Lane-Getaz, 2012; Delmas et al., 2007).

Two questions remain: How scientific (reliable and valid) is the available evidence, and to what extent is it generalizable? Becker (1996) published a quasi meta-analysis but with minimal reference to reform and introductory statistics, and recent systematic reviews (Tishkovskaya & Lancaster, 2012; Zieffler et al., 2011) have not specifically addressed the quality of the research evidence.

OBJECTIVE

This paper presents the results of a qualitative systematic review of the literature on introductory statistics education (at the college level) so as to better understand the status of reform efforts, regarding course content, pedagogy, assessment, and the integration of technology. In addition to gaining insight into the effectiveness of reform strategies, such information facilitates evidence-based instruction, and informs policy and decision-making regarding the needs of the discipline, including resource allocation, and research priorities.

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

METHODOLOGY

A qualitative systematic review limited to research articles published in the *Journal of Statistics Education* (JSE, n = 73, period reviewed: 1993 – 2013), *Statistics Education Research Journal* (SERJ, n = 26, 2002 – 2013), *Technology Innovations in Statistics Education* (TISE, n = 19, 2007 – 2013), and *Numeracy* (n = 10, 2008 – 2013) was conducted. These journals are widely recognized as the major peer-reviewed publication outlets in the discipline, particularly with reference to introductory statistics (at the college level) and reform, aimed at fostering statistical literacy; which themes constituted the inclusion criteria of this study. Also, pioneer educators were engaged via email to help to identify key articles. The papers were evaluated based on the quality, quantity, and consistency of the evidence regarding course content, pedagogy, assessment and integration of technology. The evaluation was performed by a statistics educator and researcher in conjunction with another faculty member involved in the teaching of quantitative reasoning. Thematic analysis techniques were used. Consistent with best practices for systematic reviews, a quality score was derived to rate the evidence for each reform domain (Table 1). The generalizability (or external validity) of the evidence was rated as *NA* (no meaningful assessment is possible, given the limited and poor quality evidence), *Low*, *Moderate*, or *High*.

Table 1: Research Evidence Rating Scale

Score	Explanation
1	Significantly lacking or non-existent; there is little or no scientific evidence, or mostly qualitative reports; with no meaningful basis for generalization
2	Emerging, meaningful, and coherent; the evidence is derived mostly from scientific research methodologies, and allows for generalizability of selected strategies
3	Substantive, convincing, and generalizable; includes evidence from experimental studies

RESULTS

Table 2 is a summary of the results of the qualitative systematic review of published literature on introductory statistics. Note that some of the articles are not exclusive to this focus. In general, the quality of the research data is poor, and evidence of the effectiveness of reform-based instruction is minimal or non-existent. There is little scientific evidence, or mostly qualitative reports; with no meaningful basis for generalization. The exception is cooperative learning (an active pedagogical strategy) for which there is substantive scientific evidence to reasonably allow for wider application, given its effectiveness in facilitating statistical literacy. Also, the evidence on authentic assessment is promising, but its effectiveness is largely limited to favorable changes in student attitudes toward statistics (rather than academic learning outcomes).

CONCLUSION & IMPLICATIONS

This qualitative systematic review of published literature on introductory statistics education and reform was limited to research articles in four English Language peer-reviewed journals, and therefore, publication bias must be considered, as well as subjectivity bias associated with the qualitative evaluation approach, including the use of quality scores. The review focused on course content, pedagogy, assessment, and integration of technology. After almost two decades of reform efforts in introductory statistics education, there is a lack of a coherent body of research, and, in general, there is no clear direction as to the effectiveness of reform strategies. Indeed, there is no recognized consensus as to indicators and measures of effectiveness. Two exceptions are, cooperative learning, and authentic assessments, for which there is a substantive amount of evidence supporting their effectiveness in facilitating positive change in statistical literacy and attitude, respectively.

Cooperative learning, also referred to as collaborative learning, team-based learning or small-group learning, is clearly the most evidence-based reform strategy, directly linked to favorable academic learning outcomes. This is unlike authentic assessments, which have been shown to be effective mostly in relation to student attitude. Indeed, attitude change is necessary, but not sufficient as a measure of the effectiveness of reform-based instruction. Moreover, the literature

is replete with poor quality research designs, case studies, and heterogeneity in research conceptualization and focus, which together, do not allow for meaningful synthesis of the research findings. Evidence-based instruction in statistics education is lacking, and for the discipline to advance and be accepted, we must urgently heed a call for change toward a more organized and scientific approach to measurement, research, and setting priorities. The mantra going forward should be: *show me the data*.

Table 2: Summary of a Systematic Review of Literature on Introductory Statistics Education Journals: *JSE, SERJ, TISE,* and *Numeracy*

Reform Domain	**Data Quality Score**	**External Validity (Generalizability of Research Evidence)**	**Comment**
Course Content (including sequencing of topics)	**1**	*NA* (Based on limited evidence, **N**o meaningful **A**ssessment is possible.)	There is a dearth of research on what topics (and sequence) should constitute an introductory course. The articles are mainly case studies addressing mostly probability, sampling, and distributions.
Pedagogy (active learning strategies)	**2** (This is based on the evidence supporting cooperative learning, otherwise, a score of **1** applies.)	*Moderate to High* (This applies to cooperative learning.)	There is a large body of information on active learning strategies, with broad assumptions about their appropriateness and effectiveness. However, empirical evidence is lacking; *except for cooperative learning (Appendix I),* for which there is substantive supporting evidence, including a few experimental/quasi-experimental studies.
Assessment of Student Learning	**1** (Academic learning outcomes) **2** (Non-cognitive outcomes, particularly attitude)	*Low* (For academic learning outcomes) *Moderate* (For non-cognitive outcomes, particularly attitude)	There is a substantive amount of information on authentic assessment, but mostly linked to student attitude rather than academic learning outcomes; scientific evaluation is lacking. Also, the literature reflects the development of psychometric instruments necessary for effective assessment.
Use and Integration of Technology	**1**	*Low*	There is much information on technology, albeit largely qualitative and anecdotal, rather than empirical. Except for case studies on simulation, applets, computers, and data sets, there is a dearth of empirical evidence of effectiveness.

REFERENCES

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., et al. (2005). Guidelines for Assessment and Instruction in Statistics Education: College Report. American Statistical Association.

Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics*, 21(1), 71–90.

Cobb, G. 1992. Teaching statistics. Heeding the Call for Change: Suggestions for Curricular Action. Edited by L. A. Steen. (Washington, D.C.: Mathematical Association of America) *MAA Notes* no. 22: 3-43.

Cobb, G. W. (2013). What might a twenty-year old conference tell us about the future of our profession? *Journal of Statistics Education*, 21(2).

Delmas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28-58.

Hassad, R. A. (2011). Constructivist and Behaviorist Approaches: Development and Initial Evaluation of a Teaching Practice Scale for Introductory Statistics at the College Level. *Numeracy*, 4(2), 7.

Hassad, R. A. (2013). Faculty Attitude towards Technology-Assisted Instruction for Introductory Statistics in the Context of Educational Reform. *Technology Innovations in Statistics Education*, 7(2).

Lane-Getaz, S. J. (2013). Development of a reliable measure of students' inferential reasoning ability. *Statistics Education Research Journal* (SERJ), 12(1), 20-47.

Pearl, D.K., Garfield, J.B., delMas, R., Groth, R.E., Kaplan, J.J., McGowan, H., and Lee, H.S. (2012). Connecting Research to Practice in a Culture of Assessment for Introductory College-level Statistics.

Rossman, A. (2011). Interview with Joan Garfield. *Journal of Statistics Education, 19*(3).

Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: a review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2).

Zieffler, A., Garfield, J., delMas, R. C., Le, L., Isaak, R., Bjornsdottir, A., & Park, J. (2011). Publishing in SERJ: An analysis of papers from 2002–2009. *Statistics Education Research Journal*, 10(2), 5-26.

Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The Statistics Teaching Inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, 20(1).

Appendix I: Selected Research Articles in Support of Cooperative Learning

Bailey, B., Spence, D. J., & Sinn, R. (2013). Implementation of Discovery Projects in Statistics. *Journal of Statistics Education*, 21(3).
Clair, K. S., & Chihara, L. (2012). Team-based learning in a statistical literacy class. *Journal of Statistics Education*, 20(1).
Dingman, S. W., & Madison, B. L. (2010). Quantitative Reasoning in the Contemporary World, 1: The Course and Its Challenges: *Numeracy*: Advancing Education in Quantitative Literacy, 3(2).
Enders, F. B., & Diener-West, M. (2006). Methods of Learning in Statistical Education: A Randomized Trial of Public Health Graduate Students. *Statistics Education Research Journal*, 5(1), 5-19.
Everson, M. G., & Garfield, J. (2008). An innovative approach to teaching online statistics courses. *Technology Innovations in Statistics Education*, 2(1).
Garfield, J. (2013). Cooperative Learning Revisited: From an Instructional Method to a Way of Life. *Journal of Statistics Education*, 21(2).
Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education*, 5(3), 1.
Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, 3(2), 1-8.
Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education*, 6(3), 1-10.
Tishkovskaya, S., & Lancaster, G. (2012). Statistical education in the 21st century: a review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, 20(2), 1-55.
Verkoeijen, P. P. J. L., Imbos, T., van de Wiel, M. W. J., Berger, M. P. F., & Schmidt, H. G. (2002). Assessing knowledge structures in a constructive statistical learning environment. *Journal of Statistics Education*, 10(2).
Wenner, J. M., Baer, E. M., Manduca, C. A., Macdonald, R. H., Patterson, S., & Savina, M. (2009). The case for infusing quantitative literacy into introductory geoscience courses. *Numeracy*, 2(1), 4.
Zacharopoulou, H. (2006). Two learning activities for a large introductory statistics class. *Journal of Statistics Education*, 14(1).
Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education*, 16(2), 1-23.