Published in final edited form as:

Ann Behav Sci Med Educ. 2013 ; 19(2): 8–9.

“Raising the Bar” for Educational Science

Jason M. Satterfield, Ph.D.

University of California, San Francisco

In the current issue of *Annals*, Dr. Kristin Bell describes an evaluation and follow-up of a promising curricular intervention teaching motivational interviewing to medical students.¹ This curriculum was previously published in *JGIM*² but the current article provides a contrasting assessment tool and longer follow-up. Moreover, the current article raises a number of perennial and important issues in the realm of educational scholarship including the selection of pedagogy, designing quality educational research, and ultimately evaluating the lasting effects of educational interventions on provider behavior. By “raising the bar” in each of these essential areas, the field of behavioral science in medical education can more effectively and efficiently evolve medical providers to meet the current needs of patients and communities. Firstly, this editorial provides a summary of Bell’s curriculum and assessment tools. Next, a brief exploration of evidence-based teaching strategies and pedagogical approaches are introduced including a discussion of the “Kirkpatrick levels” commonly used for evaluations.³ Lastly, issues of educational research design quality are presented to promote ongoing quality improvement in future curricular studies.

In 2008, Bell and Cole described a 4-week, 8-hour small group curriculum teaching motivational interviewing (MI) to third-year medical students.² Teaching strategies were multi-modal and included didactics, video demonstrations, role plays, and other interactive exercises. In this pre-post, single cohort study, students completed a measure of knowledge and confidence and used a performance assessment called the VASE-R – the Video Assessment of Simulated Encounters, revised. While most prior studies relied on self-report questionnaires, Dr. Bell’s group used a validated research tool to better tap actual (simulated) performance. In the pre-post comparison, students showed significant improvements in knowledge, confidence, and MI skills.

In the current study, Dr. Bell assessed fourth-year medical students using a standardized patient as part of a Clinical Practice Exam (CPX).¹ Students who had been trained in MI in their third year (presumably the same sample for the 2008 study) were compared to a historical control group (2004–5) that had taken the same CPX but had not been trained in MI. The SP case included scores for overall satisfaction, information gathering, active listening, exploring patient perspective, addressing feelings, meeting patient needs, exploring reasons for non-adherence, and shared decision making. Since current students had been trained at various points during their third year, some students took the CPX

shortly after MI training and some (those trained at the beginning of their third year) had a delay of approximately a year between MI training and CPX testing. Overall, results showed that students trained in MI outperformed historical controls in 7 of the 8 categories. The relationship between CPX and prior VASE-R scores was not reported. Dr. Bell asserts that the SP satisfaction score is a “patient surrogate” indicating a Kirkpatrick level four evaluation – i.e., measures the effect of the intervention on real world, clinical outcomes. Time from MI training was not used as a co-variate to explain CPX performance nor were subset analyses performed to strengthen the argument of lasting behavioral changes in the learners. It is unknown if performance deteriorated over time. Drawing from both of these papers, Dr. Bell has shown the effectiveness of her curriculum using multimodal assessment tools that go above and beyond basic MI. Moreover, she raises the possibility that these effects last at least up to one year and may include effects on patient satisfaction.

While neither paper was intended solely as a curriculum development description or survey of pedagogical tools, they nonetheless raise a number of important issues that behavioral scientists and educators should consider when creating new training interventions. As there are typically far more topics than time, any curricular additions or modifications must meet a well-defined need, be evidence-based, maximize efficiency, evaluate effectiveness, and have superb “curb appeal” – often achieved by including learners and stakeholders in the curricular development process. Ideally, these new teaching interventions fit within existing infrastructures, are integrated well with broader institutional priorities and goals, and demonstrate a clear “return on investment.” Although web-only modules are inexpensive and efficient, their impact is limited as are straight didactic lectures and self-study materials. A recent systematic review identified optimal strategies for successful behavior change curricula in medical learners by abstracting content from $n=109$ peer reviewed, evidence-based curriculum studies.⁴ To meet inclusion criteria, these studies had to describe a behavior change counseling curriculum, teach medical trainees, and include an assessment of the curricular intervention. Of the studies reviewed, those with the strongest curricular (learning) outcomes employed multi-modal teaching strategies (e.g., videos, simulations, role plays, discussions) and included skills practice in simulated or actual clinical settings. Contrary to popular opinion, successful behavior change counseling curricula did not need to be overly long or time consuming with 62 studies using less than eight curricular hours and 51 studies spanning four or fewer weeks.

The outcome measures chosen in each of the Bell studies also raise important considerations when evaluating learners and behavioral curricula. In the behavioral counseling systematic review,⁴ each study outcome measure was also ranked on a standard Kirkpatrick³ level: 1) student participation or reaction (satisfaction), 2) modification of attitudes, knowledge, or skills, 3) demonstrated behavioral changes, and 4) benefits to organizational practice or patient outcomes. Bell also places her outcomes on the Kirkpatrick continuum and implies that an SP satisfaction score could qualify for the highest level 4 (changes in patient outcomes). While laudable, changes in CPX performance are typically scored as level 2 (changes in skills). A level 3 would require demonstrated changes in actual clinic behavior (e.g., via direct observation or possibly through chart reviews). A level 4 would require changes in medical or clinic outcomes (e.g., lower HgbA1c levels, fewer patients smoking).

Of the $n=109$ reviewed studies by Hauer et al.,⁴ 41 used SP exams but only 12 reached a level 4 by demonstrating actual benefit to patients or organizational practice. Simply put, evaluations of behavioral science in medical education need to move beyond simple satisfaction surveys or pre-post self-reports of skills. Bell pushes the envelope by including the VASE-R and, later, an SP exam – both examples of thoughtful and time consuming assessment tools. The hope is that Bell and her colleague may continue to climb the Kirkpatrick hierarchy to see what learners do in real clinical practice and by measuring actual patient outcomes. By achieving these “higher” Kirkpatrick levels, one hopes to capture the attention of the medical center leaders who may be persuaded to invest further resources in behavioral curricula once they see real changes in patient outcomes.

While Bell and Cole’s MI curriculum was included as a “successful” study in the aforementioned behavioral counseling systematic review, it was noted that most studies (like Bell’s) used a single cohort, pre-post design at a single institution with a modest sample size. In Cook, Beckman, and Bordage’s systematic review of the quality of reporting in medical education studies, $n=110$ studies were critically reviewed.⁵ Quality indicators included containing a critical literature review, presenting a theoretical framework, stating study intent and study design (e.g., aim, research question, hypothesis), definition of the main intervention and comparison/control group, and consideration of human subject rights. On the whole, most studies were found to be poor: 45% contained a literature review, 55% had a conceptual framework, only 16% contained an explicit study design statement, and fewer than half had a comparison group. Hauer et al.⁴ found similar quality gaps where only 23 of 109 articles used a randomized controlled design and only 43/109 mentioned IRB approval.

If the field of behavioral science in medical education is to move forward, it is essential that educators and scientists demand more of themselves and the studies they lead. The field needs larger, multisite studies with sound designs and high quality (and high impact) outcome measures. If we expect medical school deans and medical center CEO’s to invest in our educational interventions and clinical services, it is our responsibility to demonstrate their undeniable utility. While true that educational research funding is scarce and cross-school collaborations are challenging, we would better invest our time and energy in doing fewer pre-post, single cohort studies (the seeming life-blood of clinician-educator advancement) and redirecting those resources into “collaboratories” and consortia. Of course, these changes will require interventions on multiple levels. We need institutional policy changes (e.g., criteria for advancement and promotion), possible changes in department structures and support (i.e., where does behavioral science live?), new research and development funding streams, and higher quality training in educational science for the next generation of curriculum researchers. Fortunately, we have rising scholars like Dr. Bell and her colleagues and organizations like the Association for the Behavioral Sciences and Medical Education (ABSAME) who will continue to raise the bar for everyone.

References

1. Bell K. Evaluation of a behavior change curriculum: Outcomes and impact in a follow-up study. *Annals Beh Sci Med Ed*. 2013; 19:3–7.
2. Bell K, Cole BA. Improving medical students’ success in promoting behavior change: A curriculum evaluation. *JGIM*. 2008; 23:1503–1506. [PubMed: 18592322]

3. Kirkpatrick, DL.; Kirkpatrick, JD. Transferring Learning to Behavior: Using the Four Levels to Improve Performance. San Francisco, CA: Berrett-Koehler Publishers, Inc; 2005.
4. Hauer KE, Carney PA, Chang A, Satterfield J. Behavior change counseling curricula for medical trainees: A systematic review. *Acad Med.* 2012; 87:956–968. [PubMed: 22622220]
5. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: A systematic review. *Med Educ.* 2007; 41:737–745. [PubMed: 17661881]