Supporting Evidence-Informed Teaching In Biomedical and Health Professions Education through Knowledge Translation: An Inter-Disciplinary Literature Review

Rochelle E. Tractenberg [a] and Morris Gordon [b]

[a] Departments of Neurology, Biostatistics, Bioinformatics, and Biomathematics; and Rehabilitation Medicine, Georgetown University Medical Cente, Washington DC

[b] Welfare, Professionalism, Transition and Careers, University of Central Lancashire, UK

Keywords: knowledge translation, education research, systematic reviews, evidenceinformed teaching

Contact: Rochelle E. Tractenberg, Georgetown University Medical Center – Neurology, 4000 Reservoir Road NW; 291 Building D, Washington, District of Columbia 20007 United States, Email: rochelle.tractenberg@gmail.com

ABSTRACT

PHENOMENON: The purpose of "systematic" reviews/reviewers of medical and health professions educational research is to identify best practices. This qualitative paper explores the question of whether systematic reviews can support "evidence informed" teaching, and contrasts traditional systematic reviewing with a knowledge-translation approach to this objective.
APPROACH: Degrees of Freedom Analysis is used to examine the alignment of systematic review methods with educational research and the pedagogical strategies and approaches that might be considered with a decision-making framework developed to support valid assessment. This method is also used to explore how knowledge translation can be used to inform teaching and learning.
FINDINGS: The nature of educational research is not compatible with most (11/14) methods for systematic review. The inconsistency of systematic reviewing with the nature of educational research impedes both the identification and implementation of 'best-evidence' pedagogy and teaching. This is primarily because research questions that do support the purposes of review do *not* support educational decision-making. By contrast to systematic reviews of the literature, both a Degrees of Freedom Analysis (DOFA) and knowledge translation (KT) are fully compatible with informing teaching using evidence. A DOFA supports the translation of theory to a specific teaching or learning case, so could be considered a type of KT. The DOFA results in a test of alignment of decision options with relevant educational theory and KT leads to interventions in teaching or learning that can be evaluated. Examples of how to structure evaluable interventions are derived from a knowledge-translation approach that are simply not available from a systematic review.
INSIGHTS: Systematic reviewing of current empirical educational research is not suitable for deriving or supporting best practices in education. However, both "evidence-informed" and scholarly approaches to teaching *can* be supported as knowledge translation projects, which are inherently evaluable and can generate actionable evidence about whether the decision or intervention worked for students, instructors, and the institution. A Degrees of Freedom Analysis can also support evidence- and theory-informed teaching to develop an understanding of what works, why, and for whom. Thus, knowledge translation, but not systematic reviewing, can support decision-making around pedagogy (and pedagogical innovation) that can also inform
new teaching and learning initiatives; it can also point to new avenues of empirical research in education that are informed by, and can inform, theory.

PHENOMENON

The past twenty years have seen a growing emphasis within higher educational initiatives on those that have the strongest evidence base (e.g.,[1-2]). This growth – particularly in medical and health professions education[3] has been fueled by the evidence based medicine movement to support clinical decision making. This has been led by the Cochrane Collaboration, established in 1993 to*…"to promote evidence-informed health decision-making by producing high-quality, relevant, accessible systematic reviews and other synthesized research evidence"*. [4 (p.)] Output consists of review reports that are highly structured, employ transparent and reproducible methods, and are updated periodically so that the evidence base for clinical decision-making can be considered "up to date". The Campbell Collaboration (created in 2000) also established a process for evidence-informed decision-making in policy and education that is closely aligned with that of the Cochrane Collaboration. The Campbell Collaboration approach was developed specifically for the results of policy changes; it focuses on identifying evidence of whether or not the policy had achieved the ends that were stated as motivating factors for these decisions. Both initiatives define a systematic review with similar concepts: "A systematic review uses transparent procedures to find, evaluate and synthesize the results of relevant research. Procedures are explicitly defined in advance, in order to ensure that the exercise is transparent and can be replicated. This practice is also designed to minimize bias." According to the Campbell Collaboration, "A systematic review must have:

- Clear inclusion/ exclusion criteria
- An explicit search strategy
- Systematic coding and analysis of included studies
- Meta-analysis (where possible)" [5]

Also common to both initiatives is a focus on recent, empirical evidence to ensure contemporaneous conclusions about the decision of interest. This *may* include unpublished trials in progress, recent empirical studies presented in international meetings (information that is not peer reviewed including technical reports and white papers) or reference works such as textbooks.

About the same time as the creation of the Campbell Collaboration, the "best evidence medical education" (BEME) Collaboration was founded by the Association for Medical Education in Europe (AMEE). The organization's goal was to support teaching (for the medical and health professions) that uses "methods and approaches to education based on the best evidence available". [6, p. 553] AMEE has supported the development and publication of many "BEME Guides", as well as evidence informed "AMEE Guides"; both are intended to summarize the "best available evidence" to promote evidence-informed teaching.

The shared methodological attributes of Cochrane and BEME specifically are supportive of a structured and systematic search through predominantly current/recent empirical peer reviewed literature; they exclude summaries of research such as are found in reference works and textbooks. The types of research on which both Campbell and Cochrane tend to focus are typically experimental or quasi-experimental study designs executed – and peer reviewed/published – within the previous seven or fewer years. As such, these types of reviews are oriented to results that support generalizations from study data to a population, which is the case with statistical analyses generally. Since the Cochrane review methodologies focus on research with experimental design and statistical (quantitative) summarization (and Campbell as well, to some extent), both the research that is reviewed and the reviews themselves tend to be quantitative, seeking to generalize from the empirical data to populations to support general ("best evidence") conclusions. By contrast, classroom-based research, particularly in higher and professional education, tends to be less compatible with experimental or quasi-experimental designs because of the variability that each class meeting may comprise and variation in teaching and learning styles in any given interaction or course (see e.g., [7-8] Knowles et al. 2005; Diamond, 2008). Thus, evidence that can inform teaching in health professions education might incorporate qualitative results, but because of the structure of the search and screening processes in "systematic" reviewing, qualitative results would not be utilized or strongly weighted in a "systematic review" of the literature in the domain. Moreover, experimental designs tend to require that context (e.g., classroom, point in curriculum, and program) and participants (learners and instructors) be fixed to the extent possible. However, this is rarely either feasible or desirable in educational contexts.

The Cochrane or Cochrane-type review methodology is a positivist approach to generalizing quantitative data from individual studies to a population. An alternative approach to utilizing educational research to inform teaching is the generalization of the data from individual studies to a *theory*. This is the objective in numerous qualitative methods such as case studies.[9,10] Yin (2014)[9, p. 16] defines case study research as "…an empirical inquiry that investigates a contemporary phenomenon ("the case") within its real life context, especially when the boundaries between phenomenon and context are not clearly evident." This description is appropriate for educational research, particularly when it is based on classroom teaching, learning, and their observation.

If a particular case study from an educational context with a given learner group, task, or assessment suggests that an intervention or teaching/assessment method is "successful", then evidence-informed teaching supports first the replication and then possibly the translation of this new knowledge. To replicate the phenomenon within a similar or even a different educational context –for example, an innovative method for teaching or assessing taken from the undergraduate and translated into a health professions context, it is helpful, and in some cases necessary, to incorporate theoretical justification for the innovation/change in teaching or assessment. Clarification of the intended outcome may also be required in order to more concretely define and describe the effective elements of the intervention. Theoretical justification, or other external validation, can support the development of an evaluable plan for the application of the innovation in the new context and particularly, to document the relevant outcome(s).

The incorporation of theory is a key aspect of the use of educational research evidence –whether qualitative or quantitative - to inform teaching and learning in the same or in different contexts. Unfortunately, theory is *not* implicitly (and usually not explicitly) part of the systematic review. This is less problematic in the evidence-based medicine and practice applications of the systematic review, since a great deal of theory goes into the process by which medical interventions are developed, tested, labeled, and ultimately approved –all of which are typically peer-reviewed and published. In educational research, the process by which interventions and innovations in teaching and assessment are developed, tested, and disseminated can be less procedural – making the identification of theoretical framework(s) and external justifications far more critical to support the use of peer reviewed and published evidence to inform one's own teaching and the assessment of learning that results.

In their systematic review undertaken to understand why few clinicians appear to utilize the Cochrane reviews, Wallace et al. (2012)[11] noted that "…lack of practical use of systematic reviews continues to present a major challenge to evidence uptake". (p. 12) While not referring specifically to the systematic reviews of educational research supporting medical education, this conclusion may also apply to educational research systematic reviews. Moreover, systematic reviews of educational research also tend to support the conclusion that "more research is needed", rather than providing advice or recommendations about the best way to address a given educational challenge – so that "uptake" or utilization of these results might actually be *more research*, and not actually the translation of this knowledge into teaching and learning.

It is also possible that, due to their generalization from data to populations, uptake (utilization, incorporation) of systematic reviews of educational research is also difficult because the structure of the review typically has limited theoretical representation or consideration. As noted above, without theoretical frameworks or external validation, the replication or translation of evidence derived from educational research can be limited, or at least, its documentation in the literature could be limited since these structures support testable hypotheses about teaching, learning, and their assessment –particularly across contextual boundaries (e.g., from undergraduate to health professions). Thus, the reviews, their context, and their theoretical grounding may all contribute to an inherently low or limited potential for contributing to the search for evidence that can inform teaching and learning.

Finally, Roberts, et al. (2015)[12] note that Cochrane reviews are based on a biased sample of research (i.e., only those that are published, which tend to favor treatments showing some effect over studies that are NOT published because of null result/no effect of treatments). Educational research is subject to the identical bias because many educational innovations are either too challenging to write up as scholarly contributions to the literature, or if no "significant" effect was observed, they will not be published or may be published in a book or compendium rather than in a peer-reviewed venue. These features of any "systematic review of the literature" may not impact the uptake of the results of such reviews, but they do limit what evidence can be used to inform both teaching *and* clinical practice if both are dependent on recent, empirical, peer-reviewed literature.

The purpose of this review is to formally explore the applicability of Cochrane-type reviewing for decision making in higher education (including medical and health professions education) and to describe knowledge translation as an alternative approach that promotes systematic summarization and utilization of evidence about teaching and learning in higher education, knowledge translation, together with a method to synthesize data collected within such reviews, the Degrees of Freedom Analysis (DOFA).

APPROACH

The relationships between educational and cognitive-psychological theories and the data that might be collected in classrooms or through higher, medical, and health professions education, are important. Cognitive and educational psychologists have longstanding, comprehensive bodies of literature and theory that relate to all adult learning (e.g., [1, 8, 12, 13, 14]) and these theoretical underpinnings support their applications in ways by instructors and in contexts that can differ from the original "empirical" findings. For clinical and biomedical data, by contrast, the clinical trials (across phases) from which data and results are systematically reviewed are driven "up the chain" by mechanistic and biomedical theory; further steps are not supportable if earlier studies and experiments are not successful. Thus, in clinical and biomedical contexts, the generalization from data (across studies systematically reviewed) to the population is supported –although limited somewhat by the fact that unpublished *negative* results are missing – brings the theoretical framework forward from earlier phase studies through to later phases and clinical trials. These empirical underpinnings for clinical and biomedical research do *not* support the application of the results in different ways and in different contexts from the specific features in which the data were generated.

A Cochrane-style summary of educational literature may characterize relevant empirical (typically recent) literature on the target topic, but it is not oriented to decision-making for educators. An alternative approach that involves an ontological shift to applying evidence to teaching and learning is "knowledge translation". In this approach, the vast educational, cognitive, and psychological literature of what is known about teaching and learning in adults –in and outside of professional training contexts (e.g., [1, 8, 12, 13, 14]) – can be translated into defensible and evaluable innovations in health professions education. A method for synthesizing literature that can be used to support the translation of knowledge obtained from empirical results or theory directly to the teaching or assessment of any instructor is the degrees of freedom analysis (DOFA). [15] The DOFA functions to generalize, or test the compatibility of case data to a specific theory.[15,16, 17] If educators are considering a change to a particular teaching or assessment method, rather than consulting the systematic reviews on that method to obtain an expected effect size or estimate of the amount and quality of empirical literature on the method, they could instead synthesize empirical and/or theoretical literature discovered systematically within the DOFA framework (see Tractenberg in review). [16] The resulting DOFA would

comprise experimental, theoretical, and/or qualitative educational research, papers, or books addressing the single, well-defined question of making a change to their pedagogy or practice. That is, the DOFA framework is constructed around a decision, not a research question. Each paper or book would be a single case and the evidence that any given "case" provides about the theory underlying the question of pedagogy or practice can then be extracted. The creation of this DOFA represents a systematic collection of evidence to support a decision that would translate this knowledge –or not – into a new context, pedagogy, or practice. Evaluation of the DOFA represents the synthesis of this systematically reviewed evidence. The outcome is a translation of the knowledge, according to how closely it is aligned with theory, into a defensible and evaluable decision. The quality of theory or findings that are incorporated into the DOFA can be explicitly included within the DOFA through point allocation (i.e., instead of assigning one "point" for "well aligned" and zero points for "unaligned", one-half point can be assigned to "marginally aligned", or other point systems can be devised).

As an example, consider the choice of whether to integrate peer-to-peer activities into a course that has been lecture-based since its inception. A formal systematic review approach (e.g., either Cochrane or inspired by Cochrane) would seek to address a *research* question along the lines of, "Do peer-to-peer activities result in increased retention (test scores) in first year nursing students?", while a DOFA synthesis of a review for knowledge translation would instead seek to obtain evidence to address the question, "Is adding a peer-to-peer activity going to align this course better with principles of andragogy than does the traditional lecture?". An important difference between these motivating questions is that while both questions enable assessment of the quality of the literature to be reviewed, the DOFA results do not assume that whatever test or outcome is chosen (test scores in the Cochrane review question) is both valid and reliable for detecting changes in students in response to the new activity. This is key as it renders the vast body of published works in learning theory, adult education, and cognitive psychology that do not use "classroom tests" as outcomes, eligible for consideration. Cochrane (or other formal) reviews can add theoretical frameworks in the form of a DOFA to make their results more interpretable for classroom decision making but the majority of high quality, useable evidence that exists and could actually inform knowledge translation efforts in health professions educational settings will be ignored or excluded from consideration in a Cochranetype review. This goes beyond the bias mentioned earlier (that negative results are either not written up or not published) to undermine a method that relies exclusively on recent, empirical, peer-reviewed literature.

Moreover, the form of the central question in a systematic review necessarily implies that a decision will still need to be both formulated and explored, possibly with additional research, even after the results are synthesized. However, the results of the DOFA will be interpretable – and actionable - whether or not anyone else has conducted peer-reviewed and published studies of the same question: the decision will either be supported (because it is aligned with theory or other evidence) or it will not be supported. Therefore, a systematic review, guided by a *research* and not a *teaching or assessment* question, may naturally lead to the conclusion that "more research is necessary"; this is a useful conclusion if the reader was looking for evidence to inform their research program, but it is arguably not a useful conclusion if the purpose of reading the systematic review was actually to find evidence to inform their teaching and assessment.

In this manuscript, a variety of features of reviewing methods, research types, and applications of reviews of research types are systematically identified and synthesised *using* the= DOFA method, adapted from Woodside (2010; [17] by Tractenberg, in review[16]). In a formal DOFA[17] the "results" of these tables would be summarized as counts of "hits" and "misses". In a DOFA that is adapted to support educational decision making,[16] however, the tables can also support or initiate discussion and can then also be material or evidence for future reviewers to consider as they plan their own literature reviews and educational research. The findings are each examples of DOFA that demonstrate this method in a manner that also furthers the scholarly discourse presented here.

The results of this series of DOFA alignment tables (presented in the next section) outline the compatibility of Cochrane (Table 1) or Cochrane and Campbell (Table 2) review method features with general research attributes. Each DOFA also includes consideration of evidence in teaching and assessment (to promote learning); either comparing features of different review methods with educational research features or educational decision-making that evidence would ostensibly used to inform in higher education.

FINDINGS:

In 2012, a comparison of the BEME and the Cochrane reviewing methods was published,[18] concluding that the differences between the methods are only in terms of degree (that is, that BEME reviews are executed essentially following Cochrane methodology). However, as noted above, educational research differs fundamentally from biomedical research, leading to the conclusion that the systematic reviewing of the evidence in educational research *must* also differ –fundamentally- from that done with Cochrane reviewing to support clinical decision-making. Table 1 outlines some characteristics and highlights key difference, between education research and research supporting clinical and biomedical decision-making.

INSERT TABLE 1 ABOUT HERE

Table 1 is simplified because biomedical research – like educational research – varies in methodology, quality, and rigor as well as in other characteristics of context, participants, and generalizability. The fundamental distinctions shown in

Table 1 underscore the claim that a reviewing method that works in biomedical research will not function similarly for educational research.

Haig & Dozier (2003)[19] review the varieties of sources – and their complexities – that should be utilized in searches "for medical education evidence" – but what they do not include are reference texts. This is particularly problematic because educational research has been going on for much longer than "medical education research"; therefore, an unbiased, replicable review of "research in medical education" that ignores the extant comprehensive bodies of theoretical and experimental work and syntheses based on theory plus empirical work over time, such as Fink, (2013);[13] Diamond (2009);[8] and Ambrose et al. (2010),[1] will lead to *impoverished* decisions that- while arguably based on "systematic" reviews, will not actually be based on the best possible evidence. This is one example of how differences in the characteristics for biomedical and clinical, vs. educational, research diminish the utility and functionality of literature reviews when reviewers apply methods to data with which they are incompatible.

Another dimension on which literature reviews of biomedical or clinical work (Cochrane) or political science work (Campbell) diverge from the reviews that can promote "best evidence" or evidence-based teaching and learning is the form of the question that is used to guide the evaluation of the literature that is actually reviewed. Table 2 summarizes the characteristics of literature that addresses the type and form of questions that represent evidence-based decisions from Cochrane and Campbell perspectives and for decision-making around teaching and learning.

INSERT TABLE 2 ABOUT HERE

Thus, literature reviews differ in the nature of the questions they can answer as well as the types and characteristics of the literature that each approach seeks and summarizes. Table 1 highlights some key differences between clinical/biomedical and educational research and together with Table 2 underscores the ways in which Cochrane-type reviews are not well suited for decision-making around pedagogy and practice.

While Cochrane reviews tend to follow a single pattern, there is actually a wide range of possible review types. Table 3 assesses the alignment of the fourteen review types identified by Grant and Booth (2009)[20] with three key questions about, or comprising, a decision about teaching and learning. If a review does not support decision-making about teaching, it is unlikely that teaching or assessment can be "informed" by that evidence.

INSERT TABLE 3 ABOUT HERE

Table 3 suggests that very few (3/14) of these methods for systematic review of the literature have a chance of actually making useful contributions to educational decisions to promote valid "evidence-informed teaching", as they are intended to do.

It may be the case that existing BEME and Cochrane reviews differ very little;[17] but what this implies is not that the Cochrane approach *is suitable* for best evidence informed teaching and learning in the biomedical and health professions. Instead, what it implies is that BEME reviews are actually not functioning (nor could they function) as intended. In fact, Hammick, Dornan & Steinert (2010)[21] argue that, "(a) clear *review* question lies at the heart of systematic review research". (p. 3, emphasis added) The challenge is then how to translate a clear *review* question into a decision to be made in higher education (or vice-versa) that would represent an application of the systematically-reviewed evidence to inform teaching. While "a clear *review* question" will guide the review – and reviewers – to a publishable systematic review of the targeted educational literature, it is not engineered to guide the review to an actionable educational *decision*. While the review may be *motivated* by a specific decision that a group wants to use evidence to make (e.g., "should we adopt problem-based learning?"), the typical formulation of the review question is actually quite different from –and not supportive of – applying relevant evidence to make this decision (see also [22]) or evidence that learning goals have been achieved. [22-23] The research question in a Cochrane-style systematic review tends to follow the format, "Is <intervention name> better than <alternative> to achieve <specific outcome> in/with <specific population>?" This formulation ensures that literature will be evaluable by multiple raters in a consistent and replicable way; however, such questions are essentially *incompatible* with specific decisions that the evidence identified in the review should be informing (e.g., [22-23]).

Compare the results from a formal Systematic Review of "the effectiveness of Problem- Based Learning (PBL)" (Kong et al. 2014) :[24] "PBL improves critical thinking relative to lectures if taught in two semesters but not in one; and more research with larger sample sizes is needed"(p. 459) with the results of a DOFA shown in Table 4 (adapted from [16]). The Systematic Review, which focused on the effectiveness of Problem Based Learning (PBL) on the critical thinking skills of nursing students, may suggest that a randomized controlled trial of PBL vs. lecture should be undertaken, but does not support decision-making of instructors or faculty who are actually considering integrating PBL into a course or curriculum. That is, the systematic review may support decision-making about research projects but it does not provide evidence to inform decisions about teaching and learning. The question that structures this type of review limits the applicability of its results to decision-making. Moreover, these 2014 results[24] suggest that there is insufficient evidence about PBL to support its incorporation into a course – i.e., using this the Kong et al. (2014) [24] results would likely lead to the decision to "stay with lectures".
Instead of focusing on recent empirical research, the DOFA in Table 4 aligns the alternatives under consideration by instructors contemplating incorporating PBL (decision alternatives, columns) with principles of learning (rows adapted

from Ambrose, et al. 2010). [1] Thus, Table 4 shows how to translate the knowledge from Ambrose et al. (2010)[1] into the decision-making process around whether or not to integrate PBL.

INSERT TABLE 4 ABOUT HERE.

Table 4 shows that, whether or not empirical evidence has been published recently to fully answer the review question guiding the formal systematic review, PBL can be seen to be well aligned with long-established principles of learning, as reviewed by Ambrose, et al. 2010.[1] Traditional lecturing is *not* aligned with these principles.[1] A decision to integrate PBL is therefore informed –and supported - by this theoretical evidence which is, as Ambrose et al. elaborate throughout their book, fully informed by decades of experimental results. [1] Staying with the traditional lecture format literally ignores this evidence.

INSIGHTS

Systematic reviews have not been used extensively and are not well suited (as argued above) to guide decision-making in higher education. They are tools for identifying and summarizing current empirical literature and given the centrality of a research question rather than a decision about teaching or learning or assessment to be made the types of decisions that a systematic review can best support may actually be about new directions for research rather than informing teaching or assessment with the resulting evidence. However, a key theoretical framework from educational psychology can be adapted to support decision-making about teaching and assessment in order to use evidence to support decisions about changing or retaining pedagogic practices. The approach comes from three questions developed to lead to the development of valid assessments by educational psychologist Samuel Messick:[25]

1. What are the knowledge, skills, and abilities (KSAs) the curriculum should lead to?
2. What actions/behaviors by the students will reveal these KSAs?
3. What tasks will elicit these specific actions or behaviors (that reveal KSAs)?

Developing and then articulating curricular objectives using these criteria makes the curriculum valid in this sense; describing (or contemplating) the *decisions* that should be supported once the curriculum has been successfully completed also supports decision-making around educational experiences and assessments that are being either developed or evaluated. These three questions are unquestionably important in the development or evaluation of curricula; however they can also be useful for obtaining or evaluating evidence to support decisions around teaching and learning. However, identification and evaluation or synthesis of evidence from the literature is not exactly aligned with these three questions and that may impede the perception of their utility for supporting "evidence-informed" and evaluable changes in teaching or assessment. The features of deciding to change or implement innovative teaching and learning articulated in Table 3 are actually related to the Messick questions supporting valid assessment. Applying these questions to literature – irrespective of how systematically it was collated or curated or the method used to achieve this collation/curation can explicitly support the use of evidence in teaching.

The first question, "What are the knowledge, skills, and abilities (KSAs) the curriculum should lead to?" can be restated to promote the evaluation of literature as evidence as, "Can/does the literature/source help identify options/needs for change in teaching or learning?" The previously cited books by Fink (2013) ,[13] Diamond (2009) ,[8] and Ambrose et al. (2010)[1] are examples of literature (each is actually its own comprehensive synthesis of rich bodies of literature) sources that can help instructors and course or curriculum developers to identify options for changes in teaching, assessment, or both.

The second Messick question ("What actions/behaviours by the students will reveal these KSAs?") can be restated as "Can/does the literature/source inform decisions about teaching and learning?" If an identified book, chapter, or manuscript does support a decision about teaching or learning, it would qualify as evidence that has the potential to inform teaching. Qualitative, rather than quantitative, summarization of the support for the decision (e.g., "Diamond (2009)[8] directly supports decisions at the curricular level but supports individual instructor decisions only indirectly") can be incorporated into a DOFA alignment table.

The third question, "What tasks will elicit these specific actions or behaviors (that reveal KSAs)?" is actually the most important one for promoting "evidence informed teaching" when it is restated as, "Can/does the literature/source provide evidence of whether – or how to detect that – the intended learning effects occurred?" Many reference texts (e.g., [1, 8, 13]) include specific methods for determining whether the application of the methods they describe have had the intended learning effect(s). Since their formulation in 1994, the Messick questions have facilitated an "outcomes-"[26] or "performance"[27] based approach to decision-making in education, primarily in curriculum and assessment development in higher education (see [22, 23]). With these features incorporated into the curricular structure, the program (and constituent courses and assignments) can be "…evaluated on the extent to which it had accomplished its explicit goals…". [28 (p. 2)] These questions are fundamentally different from the "research questions" that are required for systematic reviews (see e.g., [21]). By incorporating these features into systematic reviews of the education literature, decisions about educational changes that are based on the evidence in the literature can similarly be evaluated with respect to the extent to which these changes had the intended effects on learning ([22-23]).

This shifts the importance of the "systematic review" from a comprehensive (in most cases) review of recent, empirical literature to the documentation of how evidence is or could be used to inform teaching and assessment in higher

education and health professions education. That is, instead of changing decisions about pedagogy into appropriate "review" questions, a different approach is needed, that treats decisions about innovations in teaching and assessment as problems requiring knowledge translation –rather than a systematic review – of empirical literature. Knowledge translation (KT), "… primarily pertains to the assessment, review, and *utilization* of scientific research." (29; emphasis added) A formal process for KT was initially presented by Lavis et al. (2003),[29] and was discussed with respect to the challenge of changing clinical practice by Grimshaw et al. (2012) .[30] This construct is most often used to discuss navigation of the 'gap' between clinical evidence and clinical practice. In health professions education, the review, assessment, and utilization of scientific research from/about adult learning and higher education is actually stuck at "the assessment and review" stage by a near exclusive reliance on methods of systematic review deriving from the Cochrane model. A focus on *translating* this knowledge about adult learning and higher education into practice in health professions education would be a welcome change. As noted, very little in systematic reviewing (of any type) of medical and health professions educational literature supports either decision making around educational innovation or the design of effective studies of the results of such decisions. Application of the principles outlined for knowledge translation could be very helpful moving the field, and our scholarship, forward. Knowledge translation projects promote incorporation of, and focus on, the context in which the knowledge was generated; the learners and instructors; and the ways in which these can direct selection of the appropriate (to context and participants) assessment in the "new" context to which the knowledge would be translated.[1, 32] The original formulation of these questions[29, 30] was "knowledge transfer" but here we consider the features of "knowledge translation",

1. What research knowledge should be translated?
2. To whom should research knowledge be translated?
3. By whom should research knowledge be translated?
4. How should research knowledge be translated?
5. With what effect should research knowledge be translated?

Knowledge translation initiatives can support the utilization of knowledge resources (e.g., educational and cognitive psychological research and reference texts) that have not yet (or not formally) been implemented; they can also be used to structure a response to the results of a systematic review of educational research – i.e., to translate that knowledge into an intervention or pedagogic/assessment innovation whose impact can be formally evaluated for its potential to contribute evidence to the learner, the instructor, or the institution.[31, 32] Combining these five questions driving knowledge translation with the three questions for valid assessment outlined by Messick[25] can support decision-making around teaching; a new matrix to guide the development of evaluable educational interventions emerges to support knowledge translation as opportunities to study new methods for valid, evidence-informed, teaching and learning. This matrix is outlined in Table 5.

INSERT TABLE 5 ABOUT HERE

Table 5 can be easily adapted for the results of literature curated to support specific decisions about teaching and assessment in one of two ways. First, the five KT questions can be used to create individual DOFA tables that comprise the 2-3 Messick questions as columns; then the literature is summarized in rows under each Messick question. Some literature/evidence will only enter into one of these five DOFA tables, but other evidence might support more of the Messick questions or more of the KT questions. Once all of the literature is summarized, then Table 5 can be reconstituted with YES or NO or PARTIALLY (or alternatively 1, 0, and 0.5 points can be assigned to these responses, respectively) entered into each cell, to provide an overall view of the level of support for the specific knowledge translation project in mind. The second way to adapt Table 5 is to utilize either the KT *or* the Messick questions (and not both) to guide the literature evaluation. Because the Messick questions are most specifically and explicitly focused on education, these three are the recommended questions to address whenever making decisions about changing teaching or assessment practice. Thomas et al. (2014)[33] argue that "(s)ocial constructivist approaches to the science of KT have the potential to support researchers interested in examining how learning in the clinical context occurs and how new knowledge is created, disseminated, exchanged and used to inform practice." (p. 2) Since constructivism (carefully distinguished from construct*ion*ism by Thomas et al.[33]) focuses on how individuals construct and apply knowledge within the context of social situations or engagement (e.g., in classrooms rather than as self-directed learners), the alignment of this approach to knowledge translation for the development of *evaluable*[1, 32] *innovation* in higher and health professions education is both timely and relevant.

DISCUSSION

The general mis-alignment between Cochrane- and even Campbell- style systematic review (SR) methods and educational literature is one reason why it is difficult - and rare- to translate a systematic review of education into justifiable decisions about teaching and assessment. The evaluation and review of the methods for reviewing outlined by Grant & Booth (2009)[20] omit any mention of the importance of theory in the formulation of an actionable decision about teaching and learning. Similarly, in their comprehensive review of major theories that can be/are used in developing and implementing teaching strategies, Taylor & Hamdy (2013)[14] give no guidance as to how to use the evidence supporting (or deriving from) any of these theories to inform one's teaching, nor how to structure schoolrly work intended to test or document the effects of utilizing either these theories or this evidence in one's teaching. In their scoping review of the use of social constructivist learning theory in knowledge translation efforts -relating specifically to the translation of clinical or biomedical knowledge into clinical practice - Thomas et al. (2014) [33, p2 of 20] noted that

"theories have been rarely used to inform the design and evaluation of KT interventions" and although they are referring to KT interventions that bridge the clinical research-clinical practice gap, the same is true of efforts to translate knowledge from educational and cognitive psychology around teaching and learning into higher, postgraduate, and health professions education. The analyses presented here provide multidimensional explication for why the "evidence-practice" gap that is supposed to be filled by systematic reviews in their traditional (Cochrane & Campbell) modes stays resolutely unbridged. Educational research is not compatible with these approaches to systematic reviewing and this fundamental incompatibility is why evidence that is systematically reviewed tends not to be informative for teaching and learning. However, it is possible to take either original evidence about teaching and learning or systematic reviews and use the combination of Messick Validity and Knowledge Translation questions presented in Table 5 to design an instructional or]assessment intervention as classroom action research in order to actually apply this knowledge to the individual instructional opportunity or decision that is being contemplated. The documentation of effects of such decisions is also facilitated with the end results of stronger teaching and deeper learning, ([22-23]) as well as better and fuller literature that can support future evidence-informed educational decision-making and the development of actionable evidence[32] based on the translation of this knowledge to new pedagogic and/or assessment strategies in health professions education.

REFERENCES

1. Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK. *How Learning Works: Seven Research-Based Principles for Smart Teaching*. Jossey-Bass, San Francisco, 2010.
2. Clark RC. *Evidence-Based Training Methods, 2 Ed*. Association for Talent Development, Alexandria, 2014.
3. Dexter, H. and Dornan, T. Technology-enhanced learning: appraising the evidence. Medical Education 201;44:746–8.
4. Cochrane Collaborative. http://www.cochrane.org/about-us/our-vision-mission-and-principles, accessed February 18, 2015.
5. Campbell Collaborative. http://www.campbellcollaboration.org/what_is_a_systematic_review/index.php, accessed March 2, 2015.
6. Harden R, Grant J, Buckley G, Hart IR.. BEME Guide No. 1: Best Evidence Medical Education. *Medical Teacher* 1999;21;6: 553-62.
7. Knowles MS, Holton III, EF, Swanson RA. *The adult learner, 6th Ed*. Elsevier, Burlington, 2005.
8. Diamond RM. *Designing and Assessing courses and curricula: a practical guide 3rd Ed*. Jossey-Bass: San Francisco, 2008.
9. Yin RK. *Case Study Research: Design and Methods, 5th Ed.* Thousand Oaks, Sage, 2014.
10. Woodside, A. G., & Wilson, E. J. Case study research for theory-building. *Journal of Business & Industrial Marketing*, 2003;18,493–508.
11. Wallace J, Nwosu B, Clarke M. Barriers to the uptake of evidence from systematic reviews and meta-analyses: a systematic review of decision makers' perceptions. *BMJ Open* 2012;2: e001220.
12. Roberts I, Ker K, Edwards P, Beecher D, Manno D, Syndenham E. The knowledge system

underpinning healthcare is not fit for purpose and must change. *British Medical Journal*: 2015;350:h2463 (accessed online)
13. Fink LD. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses 2nd Ed.* Jossey-Bass: San Francisco, CA, 2013
14. Taylor DCM & Hamdy H. Adult learning theories: Implications for learning and teaching in medical education: AMEE Guide No. 83. *Medical Teacher* 2013;35;11: e1561-72.
15. Campbell, D. T. Degrees of freedom in the case study. *Comparative Political Studies* 1975;8: 178–93.
16. Tractenberg RE. (in review). Degrees of Freedom Analysis for Educational Research: examination of the capstone experience.
17. Woodside AG. *Case Study Research: Theory, Methods, Practice*. Bingley, Emerald Group Publishing, 2010.
18. Patrício M, Vaz Carneiro A. Systematic reviews of evidence in medical education and clinical medicine: is the nature of evidence similar? *Medical Teacher* 2012;34;6:474-82.
19. Haig A, Dozier M. BEME Guide No 3: Systematic searching for evidence in medical education - :Part 1: Sources of information. *Medical Teacher* 2003;25;4: 352-63.
20. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal* 2009;26;2:91-108.
21. Hammick M, Dornan T, Steinert Y. Conducting a best evidence systematic review. Part 1: from idea to data coding. BEME Guide no. 13. *Medical Teacher* 2010;32: 3-15.
22. Hutchings P, Kinzie J, Kuh GD. Evidence of student learning: What counts and what matters for improvement. In GD Kuh, SO Ikenberry , & NA Jankowsk (Eds.). *Using Evidence of Student Learning to Improve Higher Education* Somerset, Jossey-Bass, 2015.
23. Kuh, G. D., Ikenberry, S. O., Jankowski, N., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. *Using Evidence Of Student Learning To Improve Higher Education*. San Francisco, Jossey-Bass, 2015.
24. Messick S. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 1994;23;2: 13-23.
25. Tyler RW. *Basic Principles of Curriculum and Instruction*. Chicago, University of Chicago Press, 1949.
26. Toohey S. *Designing Courses for Higher Education*. Philadelphia, The Society for Research into Higher Education & Open University Press, 1999.
27. Marzano RJ. *Designing a new taxonomy of educational objectives*. Thousand Oaks, Corwin Press, 2001.
28. National Center for the Dissemination of Disability Research. *Focus: What is Knowledge Translation?* 2005, A Technical Brief (Number 10) from the National Center for the Dissemination of Disability Research. Accessed from http://www.ncddr.org/kt/products/focus/focus10/ on October 11, 2015.
29. Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J. How can research organizations more effectively transfer research knowledge to decision makers? *Milbank Quarterly* 2003;81:221–22.
30. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. *Implementation Science* 2012;7:50.
31. Hutchings P, Kinzie J, Kuh GD. Evidence of student learning: What counts and what matters for improvement. In GD Kuh, SO Ikenberry, & NA Jankowski (Eds.). *Using Evidence of Student Learning to Improve Higher Education*. Somerset, Jossey-Bass, 2015.
32. Hutchings, P. *Aligning educational outcomes and practices*. 2016 (Occasional Paper No. 26). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
33. Thomas A, Menon A, Boruff J, Rodriguez AM, Ahmed S. Applications of social constructivist learning theories in knowledge translation for healthcare professionals: a scoping review. *Implementation Science* 2014, 9:54

Table 1. General characteristics of the research that is targeted in the literature reviewed supporting clinical and educational "interventions"

Characteristic	**RCT/clinical/biomedical research**	**Educational research**
Primary (original) empirical research is being carried out and published currently.	**Yes** - clinical trials are appearing in peer-reviewed journals and white papers; new epidemiologic findings (supporting or suggesting clinical research) published regularly.	**No** - majority of knowledge about how adults learn, and about teaching techniques, is already in books and other not-peer-reviewed resources
Intervention is mostly standardized and response by participants is the only/main source of variability	**Yes** - majority of clinical interventions fit this description.	**No** - participants and interventions (teaching skill, fidelity to specific methodologies) both vary
Textbooks or edited volumes represent what is known about interventions.	**No.**	**Yes.**
Decision to be made about whether or not to use an intervention depends primarily on the intervention,	In general, **yes.**	In general, **no;** participant "readiness", interest, and engagement are key factors in the feasibility of an

and not on the subject/participant.		intervention.
Evidence about "effect" or "effectiveness" is primarily quantitative.	**Yes**. Meta analysis methods are appropriate for estimating the effect size for an intervention across multiple, similarly structured, studies.	**No.** Meta analysis methods are rarely appropriate
Intervention effect/effectiveness fundamentally depends on the age and/or sex of the participant/subject.	**Yes.** Interventions are most often administered to individuals, and dosing can depend on sex, size, and other developmental (e.g., age, family medical history) considerations.	**No**. Interventions – usually administered to groups rather than individuals, typically differ based on the *intention* with which they are chosen, rather than based on participant characteristics like sex and size.

Table 2. Features of the literature review within methods used in Cochrane and Campbell Collaboratives, vs. reviews for knowledge translation to support educational decisionmaking.

Feature:	Cochrane SR	Campbell SR	Knowledge Translation for decision-making around pedagogy and practice.
Question form: Characteristics:	Does treatment X for condition Y have an effect that is clinically or statistically better than effect/treatment Z (or placebo) for condition Y?	Does policy/intervention A achieve its objective in terms of outcomes B, C, or D; and what are the barriers or facilitators of intervention A's effectiveness and sustainability?	Does a particular method for teaching or assessment lead to better rates of achievement of specific learning objectives than my current method(s)? Is learning more cognitively complex or sustainable with another method?
Focused mainly or solely on empirical (peer-reviewed & published) research	Yes	Yes	No

Reviews empirical (relatively current) research that is based on a trajectory of peer reviewed, published work	Yes (phase I, II and III clinical research)	Partly (no specific trajectory but historical case analyses and contextual reviews)	No
Materials reviewed includes reference materials (e.g., books)	No	No	Yes
Accommodates qualitative as well as quantitative results	No; qualitative reviews are carried out separately (on questions with not effectiveness outcomes).	Partly	Yes
Inclusion and exclusion criteria for including a resource/references are clear	Yes	Yes	Partly
Explicit search strategy	Yes	Yes	Not always part of the protocol but **CAN be included.**
Systematic coding and analysis of included studies	Yes	Yes	Yes
Review seeks to generalize the reviewed data to the population	Yes	Yes	No

Table 3. Systematic review types and their potential to contribute (yes/no/partially) to decisions about teaching and learning.

Decision features about teaching and learning: Review Type[20]:	Can/does it help identify **options/needs for change** in teaching or learning?	Can/does it help **inform the/a decision** about teaching or learning?	Can/does it **provide evidence of whether the intended learning effects occurred**?
Critical Review ("presents, analyses and synthesizes material from diverse sources")	PARTIALLY (p)	p	p – diverse sources includes books and other grey materials
Literature/Narrative Review ("reviews published literature… possibly peer-reviewed")	p	p	p – if going beyond current empirical research, y
Mapping Review/ systematic map ("map out and categorize existing literature on a	YES (y)	NO (n)	n

particular topic… identifying gaps")			
Meta-Analysis * ("statistically combines the results of quantitative studies to provide a more precise" estimate of an effect)	N/A	N/A	N/A
Mixed studies/methods review ("most frequently refers to bringing together…a quantitative effectiveness review and a qualitative review on attitudes")	p	p	p
Overview ("any summary…that attempts to survey the literature and describe its characteristics")	n	n	n
Qualitative systematic review/evidence synthesis ** (integrating or comparing the findings from qualitative studies." Also called "qualitative evidence synthesis")	n	p	p
Rapid Review/Rapid Evidence Assessment ("assessment of what is already known about a policy or practice issue.")	n	n	n
Scoping Review ("a preliminary assessment…to identify the nature and extent of research evidence")	n	n	n
State-of-the-art Review ("subtype of the more generic 'literature review'…tend to address more current" research)	n	n	n
Systematic Review ("systematically search for, appraise and synthesize research evidence.")	n	n	n
Systematic Search and Review *** ("combines …a critical review with a comprehensive search…the result is a 'best evidence synthesis'.")	n	n	p
Systematized Review ("include one or more elements of the systematic review while stopping short of claiming that the …output is a systematic review.")	n	n	n
Umbrella Review ("a mechanism for aggregating findings from several reviews that address specific questions")	n	n	n
* In their paper, Grant & Booth (2009)[20] separate meta analysis, the typical result in a			

Cochrane review, and other types of systematic review, including the subtype, “systematic review”.
** **this table** is an example of a qualitative evidence synthesis.
*** this is actually what BEME reviews are intended to be.

Table 4. Degrees of Freedom Analysis: Alignment of two decision alternatives, integrating PBL, vs continuing with traditional lecture, with principles of learning.

	INTEGRATE PBL §	CONTINUE WITH LECTURES *
Prior knowledge can be helpful	YES - PBL based in part on this principle	No
Knowledge organization supports learning and application of new knowledge	PARTIALLY – PBL uses authentic problems to frame and motivate learning; whether it is sustained is unknown.	No
Promotes motivation to learn/sustain learning	PARTIALLY – PBL uses authentic problems to frame and motivate learning; whether it is sustained is unknown.	No
Mastery is supported (opportunities to acquire component skills, practice integration, and learn when to apply them)	PARTIALLY – only full integration of PBL throughout a curriculum, with explicit developmental trajectories for the target knowledge, skills and	No

	abilities, will promote mastery	
Goal-directed practice with formative feedback provided	PARTIALLY- PBL is essentially goal-directed learning and practice; formative feedback may *or may not* be included.	No
Course climate supports learning	YES -PBL is an authentic application of critical thinking.	NO
Students will learn to monitor and adjust their approaches to learning	YES if well integrated – if PBL is integrated one time into one course in an entire curriculum, then NO.	NO - if courses in the curriculum vary, students will learn to adjust; this will be contingent on assessment strategies across the curriculum and not this specific lecturebased course.
§ Assumes that PBL is fully integrated into this course –with introduction and practice throughout the course - not simply used in a single class meeting. * Assumes that students passively attend lecture without interactivity; notes may (students do not actually need to attend) or may not (students must attend to get notes) be made available.		

Table 5. Alignment of Messick Validity and Knowledge Translation questions for decisions to implement new methods for teaching and learning.

Messick Validity: **Knowledge** **Translation:**	What are the knowledge, skills, and abilities (KSAs) the curriculum should lead to?	What actions/behaviours by the students will reveal these KSAs?	What tasks will elicit these specific actions or behaviours (that reveal KSAs)?
What research **knowledge** should be translated?	Knowledge to be translated should specify the KSAs and how the new teaching or assessment method furthers the achievement of the curricular learning goals.	Translation of target knowledge should have specific behaviours/actions that students *can* exhibit within the learning context to indicate whether learning occurred.	The difference(s) between the existing and new method of teaching or learning should be clear from the tasks that students are asked to perform.
To whom should research knowledge be translated?	At what level, and/or at what point in the curriculum, is the new educational research	Students participating in the new method must be able to benefit/learn from it, and this learning must be detectible. Determining what students would do to demonstrate that they	

	knowledge most productively translated?	have learned as a result of the translated knowledge	
By whom should research knowledge be translated?	Is there any specific training that instructors need in order to develop/maintain fidelity to a new method? Is that training in place/available?	There may need to be multiple assessments to triangulate effects of teaching on learning when students have multiple avenues to completing tasks/executing behaviours that are targeted.	
How should research knowledge be translated?	Is the new method an application of an established method in an innovative way, or a new way of applying an established method?	It is critical that knowledge/research about teaching that is hypothesized to have an impact on learning is studied or introduced within a structure where the teaching and the assessment are aligned.	
With **what effect** should research knowledge be translated?	What is the effect the KT will have? Will it generate clearer evidence that students have learned? Will it create better alignment of courses with curriculum objectives?	The ways in which the educational research/knowledge that is to be translated will affect learner behaviours or actions must be contemplated before an intervention is initiated; otherwise, no data can be collected on the efficacy or effect of the translation effort *on learning*.	The effect/efficacy of translated knowledge must be detectible in student behaviors/tasks, and these tasks must be aligned with the instruction to support conclusions about the effect of the intervention on learning or other relevant outcomes.