

ACADEMIC MEDICINE

Journal of the Association of American Medical Colleges

Uncomposed, edited manuscript published online ahead of print.

This published ahead-of-print manuscript is not the final version of this article, but it may be cited and shared publicly.

Author: Rachel H. Ellaway, PhD; Bridget C. O'Brien, PhD; Jonathan Sherbino, MD, Med; Lauren A. Maggio, PhD; Anthony R. Artino Jr, PhD; Laura Nimmon, PhD; Yoon Soo Park, PhD; Meredith Young, PhD; Aliko Thomas, PhD, OT

Title: Is There a Problem With Evidence in Health Professions Education?

DOI: 10.1097/ACM.00000000000005730

ACCEPTED

Academic Medicine

DOI: 10.1097/ACM.0000000000005730

Is There a Problem With Evidence in Health Professions Education?

Rachel H. Ellaway, PhD¹, Bridget C. O'Brien, PhD², Jonathan Sherbino, MD, Med³, Lauren A. Maggio, PhD⁴, Anthony R. Artino Jr, PhD⁵, Laura Nimmon, PhD⁶, Yoon Soo Park, PhD⁷, Meredith Young, PhD⁸, and Aliko Thomas, PhD, OT⁹

¹**R.H. Ellaway** is professor, Department of Community Health Sciences, and director, Office of Health and Medical Education Scholarship, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; ORCID: orcid.org/0000-0002-3759-6624.

²**B.C. O'Brien** is professor, Department of Medicine, and education scientist, Center for Faculty Educators, University of California, San Francisco, San Francisco, California; ORCID: orcid.org/0000-0001-9591-5243.

³**J. Sherbino** is professor, Department of Medicine, and assistant dean of health professions education research, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada.

⁴**L.A. Maggio** is professor, Department of Medicine and Health Professions Education, Uniformed Services University, Bethesda, Maryland; ORCID: orcid.org/0000-0002-2997-6133.

⁵**A.R. Artino Jr** is professor, Department of Health, Human Function, and Rehabilitation Sciences, and associate dean for educational research, School of Medicine and Health Sciences, George Washington University, Washington, DC; ORCID: orcid.org/0000-0003-2661-7853.

⁶**L. Nimmon** is scientist, Centre for Health Education Scholarship, and associate professor, Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, British Columbia, Canada; ORCID: orcid.org/0000-0002-7291-603X.

⁷**Y.S. Park** is Ilene B. Harris Endowed Professor and head, Department of Medical Education, University of Illinois College of Medicine, Chicago, Illinois; ORCID: orcid.org/0000-0001-8583-4335.

⁸**M. Young** is associate professor, Institute of Health Sciences Education, Faculty of Medicine and Health Sciences, McGill University, Montreal, Quebec, Canada; ORCID: orcid.org/0000-0002-2036-2119.

⁹**A. Thomas** is associate professor, School of Physical and Occupational Therapy and Institute of Health Sciences Education, Faculty of Medicine and Health Sciences, McGill University, Montreal, Quebec, Canada; ORCID: orcid.org/0000-0001-9807-6609.

Correspondence should be addressed to Rachel H. Ellaway, Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3280 Hospital Dr. NW, Calgary, Alberta, T2N 4Z6 Canada; email: rachel.ellaway@ucalgary.ca.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Disclaimers: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, the Department of Defense, or the U.S. Government.

ACCEPTED

Abstract

What constitutes evidence, what value evidence has, and how the needs of knowledge producers and those who consume the knowledge produced as evidence might be better aligned are questions that continue to challenge the health sciences. In health professions education (HPE), debates on these questions have ebbed and flowed with little sense of resolution or progress. In this article, the authors explore whether there is a problem with evidence in HPE using thought experiments anchored in Argyris' learning loops framework.

From a single-loop perspective ("How are we doing?"), there may be many problems with evidence in HPE, but little is known about how research evidence is being used in practice and policy. A double-loop perspective ("Could we do better?") suggests expectations of knowledge producers and knowledge consumers might be too high, which suggests more systemwide approaches to evidence-informed practice in HPE are needed. A triple-loop perspective ("Are we asking the right questions?") highlights misalignments between the dynamics of research and decision-making, such that scholarly inquiry may be better approached as a way of advancing broader conversations, rather than contributing to specific decision-making processes.

The authors ask knowledge producers and consumers to be more attentive to the translation from knowledge to evidence. They also argue for more systematic tracking and audit of how research knowledge is used as evidence. Given that research does not always have to serve practical purposes or address the problems of a particular program or institution, the

relationship between knowledge and evidence should be understood in terms of changing conversations, as well as influencing decisions.

ACCEPTED

“I believe in evidence.” – Isaac Asimov, *The Roving Mind*¹

This is an age seemingly awash with evidence, with many discussions regarding its nature, provenance, utility, and strength. Debate continues as to what kinds of evidence should (and should not) guide policy and practice. Much of this activity is constrained by a lack of clarity regarding when and how knowledge becomes evidence. The evidence-based medicine movement, which promoted the use of evidence in decision and policy making, has provided theories of evidence,² but these have tended to be more focused more on provenance (i.e., how evidence is produced) than on utility (i.e., value in practical applications).³ Although evidence-based medicine has been the target of significant criticism among philosophers of evidence, particularly regarding the constraints it places on methodological rigor and how that limits other ways of approaching problems in clinical care and public policy,⁴ issues of what constitutes evidence and what value evidence has remain.

What Is Evidence?

In making these opening statements, we acknowledge that the terms information, knowledge, and evidence are often used interchangeably, and that they hold different meanings in different contexts and disciplines. In this article, we use the term *knowledge* to mean sources that convey authority and rigor, such as reports on research studies that were conducted using rigorous methods and articles expressing logically and well-grounded expert opinions. While there are other sources of knowledge beyond those published in peer-reviewed journals, they are not the focus of this article. We focus our discussion on the knowledge generated by

scholarly activity and published in peer-reviewed journals, and how this knowledge becomes *evidence*.

When it comes to knowledge or evidence, distinctions have been made based on roles. Here, we use Albert et al.'s framing of knowledge producers and knowledge consumers,⁵ such that *knowledge producers* refers to researchers and scholars who create knowledge of various kinds, and *knowledge consumers* refers to leaders, teachers, instructional designers, and policymakers (among others) who use this knowledge to guide decisions, create or change policy, inform actions, and support decision-making. Although we use these terms in a somewhat dialectical fashion to advance our arguments, we fully acknowledge that the distinction between them is often unclear, given that many individuals perform both roles at different times.

Evidence is a relative concept that has been described both in terms of how it is produced (i.e., what kind of research or scholarship generated it) and in terms of its use, including the contexts in which it is considered and negotiated.⁶ However, provenance (e.g., that it was generated through research) is not, we would argue, the constitutive characteristic of evidence. Research and other kinds of scholarly activity, such as program evaluations and knowledge syntheses, may produce or aggregate knowledge, but we argue that knowledge only becomes evidence when it is used to inform or support a particular position or claim. Given the many debates regarding what evidence is, what it means, and what it does or is meant to do, it seems that the term *evidence* may create more problems and raise more questions than it is likely able to resolve.

Health professions educators have also grappled with questions of evidence, including what constitutes evidence, what uses it can or should be put to, and how its quality or reliability might be appraised.⁷⁻¹² Indeed, given the volume of work published in health professions education (HPE) and the apparent enthusiasm for evidence-informed practice,^{7,11,13} one might expect the HPE evidence base to be robust and useful, based on a consensus as to what constitutes good evidence.¹⁴ In reality, there are longstanding concerns that the evidence base in HPE has had little impact on practice or on the quality of educational practice,^{14,15} that the evidence base is relatively weak,¹⁶ that it is believed to lack relevance to educational practitioners,¹⁷⁻¹⁹ that it is too detached from the complexities of context,^{15,20} and that it is hard to translate into practice.²¹ It has also been suggested that knowledge from research is often not articulated in ways that resonate with decision-makers and other knowledge consumers (e.g., teachers, program directors), which can also diminish its utility as evidence in driving change.²² Take, for example, an HPE instructor who seeks to use the findings of a research study in their teaching, but falters when confronted by curriculum constraints, accreditation standards, a lack of resources and time, and a paucity of participation in a given practice change or innovation by colleagues and learners. Knowledge from research that fails to traverse this challenging (but all too familiar) gap from publication to application in local contexts may have little real impact on what that instructor can do.

We do not want to join the queue of scholars wringing their hands and bemoaning the current state of HPE research knowledge that is used as evidence. Nor do we wish to suggest ever more inventive ways to close presumed evidence-to-practice gaps. There are many who have tried to

do so, and their efforts seem to have had as little impact on educational practice as the research they sought to mobilize.^{23,24} Instead, we aim to reconceptualize the problem with evidence in HPE by using thought experiments framed by Argyris' loop learning framework²⁵: We ask whether there is a problem with evidence in HPE; if there is a problem, then what can be done to resolve it; and are HPE researchers and health professions educators asking the right questions when it comes to evidence in HPE? Answers to these questions can provide a foundation for discussions between knowledge producers and consumers to identify common scholarship goals and plan scholarly activities that better balance the needs of scholars, educators, and leaders. More broadly, we seek to guide a balanced and thoughtful discussion of the role of evidence in HPE, the responsibilities of both knowledge producers and consumers when engaging with evidence, and how HPE as a field can anchor a reasoned expectation of what evidence can do, should do, and will not be able to do.

Single-, Double-, and Triple-Loop Learning

We use the framing of single-, double-, and triple-loop learning^{25,26} to ask whether there is a problem with evidence in HPE. In Argyris' framework, single-loop learning ("How are we doing") explores whether a given activity was executed as planned. Double-loop learning ("Could we do better?") explores what might be done about any problems or shortcomings that are identified in executing an activity. Triple-loop learning ("Are we asking the right questions?") explores whether the activity was even necessary or whether its goals might be achieved in some other way. Learning takes place at the end of each loop, where those responsible for the activity examine what happened and reflect on how the next iteration of the activity might be

approached. These loops are analytical and are not intrinsically hierarchical or mutually exclusive. Different loops can focus on different actors, attending to individuals for whom the activity is important or who are struggling with the activity. In our framing, in loop 1, we focused on knowledge consumers and their efforts to translate knowledge producers' work into practice. In loop 2, we focused on interactions between knowledge producers and consumers, and in loop 3, we focused on knowledge producers' roles and responsibilities in the use of knowledge as evidence.

We recognized that there are 2 potential evidence-to-practice gaps that are frequently discussed in the HPE literature. First, there are clinical evidence-to-practice gaps that HPE may seek to address (for instance, through continuing medical education),²⁷ where educational modalities are used to implement clinical practice changes based on evidence. Second, there are educational evidence-to-practice gaps that are about the training itself.²⁸ These latter gaps are our focus in this article.

Loop 1: Is there a problem with evidence in HPE?

Some have argued that there are no major problems with HPE evidence,²⁹ while others have identified many problems.^{11,12} How we answer our loop 1 question, therefore, depends on what we mean by evidence.

Starting with the conceptualization of evidence as knowledge generated through rigorous research (provenance), we would observe that funding agencies and journals are the

gatekeepers, deciding what research is scientifically sound and worthy of funding or publication. Although we might, from this perspective, argue that there is nothing intrinsically awry with the production of research evidence, the use of markers of research quality is notoriously beset with problems, not least of which is the irreducibility of the quality of social science research to checklists and algorithms.³⁰ Moreover, it still seems that educators are, as a whole, not well-engaged (and sometimes not engaged at all) with the research that has been produced ostensibly for them.^{12,21,31,32} Not only is provenance no guarantee of application, understanding why this disconnect between rigorous knowledge production and application persists can be challenging.⁵ For instance, some journals emphasize the importance of discussing the practical relevance or implications of a study's findings, suggesting that authors of peer-reviewed publications bear much of the responsibility for the translation of their work. An alternative position is that relevance involves an inevitable negotiation of the practical needs of knowledge consumers with the theoretical richness and methodological rigor of knowledge producers, which in turn could challenge the validity of producers' claims about the relevance and value of the knowledge they draw on. Certain knowledge may have been relevant or useful in knowledge producers' own contexts, but that does not mean that it generalizes or transfers to other contexts.²⁰ A lack of standardization in both education and research practices and a lack of attention to the role of context can also undermine the utility of the findings of published studies.³³ Simply put, although there are mechanisms for ensuring that the knowledge produced is of good quality, verification of research quality does not guarantee that knowledge consumers will use the knowledge, and if the knowledge is not used, then it fails to become evidence.

The conceptualization of evidence as reflected by its utility suggests that the primary role of HPE research is to produce evidence to guide practice and/or policy. Whether this is or should be the case has been fiercely debated.^{20,34} Given that knowledge consumers, such as decision-makers, draw on multiple knowledge sources, such as personal experiences and preferences or peer recommendations,³⁵ scientific evidence (i.e., systematic, methodical, transparent, and made credible through peer-review) is only one source of evidence among many and may not necessarily be the most important or useful. Moreover, the existence of scientific knowledge does not mean it will be found and appraised, let alone used, as evidence, particularly if other sources of knowledge, such as personal beliefs and experiences, are more convenient or politically expedient. In this conceptualization, there are relatively weak and circumstantial connections between scientific knowledge and its use. This is not necessarily a problem given that, while a particular article may not be used to directly inform a particular decision, it may still influence the decision by shaping the broader discourse on the issue at hand.

Consider, as an example of this broader influence, Lucey et al.'s description of how thinking about equity and fairness in U.S. medical schools has shifted in the face of research findings describing continuing disadvantages to patients, a continuing lack of diversity in U.S. medical school classes, and ongoing systemic bias in assessment.³⁶ Another example is how conversations about fairness in peer review have shifted based on research demonstrating low publication rates of articles submitted from countries in the global south.³⁷ Indeed knowledge from research may have practical implications or be influential in many ways, such as changing the kinds of questions knowledge consumers ask, the kinds of knowledge they consider as

evidence, and the kinds of solutions they pursue. It may also be that a study's relevance or potential to be used as evidence is perceived differently by knowledge producers and consumers, reflecting an intrinsic subjectivity in appraising the utility of any given study. Although measures of study quality can help identify issues of rigor and reliability,³⁸ the ultimate arbiter of moving knowledge into evidence, based on its utility, is in the eye of the knowledge consumer.

In this first loop, which asks whether there is a problem with evidence in HPE, there may indeed be many problems with evidence (although the extent and breadth varies according to different interpretations of what evidence is or should be). This uncertainty is exacerbated by not knowing whether, how, or when research knowledge is being used to inform HPE practice.^{28,39,40} The data currently being collected, such as citations, downloads, and media mentions, is at best loosely correlated with what educators are doing as educators. Different metrics and methods are required to ascertain to what extent and in what ways knowledge is being used as evidence. Moreover, this loop suggests that knowledge producers should be less focused on citations as indicators of impact, and that knowledge consumers should be more aware of how evidence influences their approach to problems and not just how they solve them. For example, the United Kingdom's Research Excellence Framework defined impact as "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia."⁴¹ To demonstrate evidence of such impact, higher education institutions were required to provide impact case studies that described research impacts and corroborating evidence of those impacts, such as media and

social media outputs, website hits and downloads, resource uptake percentages, workshop attendance, learner or organizational outcomes, and practice and policy changes. This way of thinking about and documenting how evidence is used might be a more robust way of evaluating whether there is a general problem with evidence in HPE.

These limitations notwithstanding, we move to loop 2, which lets us ask what might be done about the problems that we have provisionally identified.

Loop 2: If there is a problem, what can be done to resolve it?

Although different kinds of knowledge may be used as evidence, for the sake of parsimony, we focus on the use of formal scientific knowledge as evidence. In doing so, we consider knowledge producer–side problems, knowledge consumer–side problems, and misalignments between the goals of knowledge producers and consumers. Certainly, there have been criticisms that much of HPE research is conducted by individuals and small teams who pursue small-scale studies exploring topics reflecting local needs, opportunities, and personal interests.^{42,43} There have also been criticisms that research is too practical,^{44,45} that it is too theoretical,⁴⁶ and that the theoretical and practical are often in tension.²⁰ Others have criticized the dominance of anglophone, Western, and northern hemisphere perspectives in medical education scholarship.¹⁴

First, let us consider knowledge producer–side problems. Issues of knowledge relevance can be addressed by researchers asking questions that matter to knowledge consumers (e.g., teachers,

program directors, policymakers), and being more strategic, rigorous, and attentive to context and the meaningful implications of findings.^{47,48} However, there is a limit to how much of this can be laid at the door of individual knowledge producers who are themselves constrained by time, resources, ethics, and policy. They are also limited by which problems are considered priorities, which problems are considered relevant to the field, and appropriate ways to approach these problems. Solutions to such knowledge producer–side problems might involve minimizing these constraints and working together as a field to collaboratively set research priorities that are sensitive to the needs of knowledge consumers.

One example of priority-setting strategies is funding agencies using grants to direct knowledge producers' attention toward challenges associated with implementing major policy changes in medical education, such as those associated with competency-based medical education.⁴⁹

Another strategy is to ensure that research teams include knowledge consumers, such as teachers, learners, or policymakers.²² A third strategy might involve HPE funding agencies setting priorities in partnership with both knowledge producers and consumers.⁵⁰ These examples notwithstanding, there is little systemic strategic direction for how to identify and solve what are collectively decided to be the priority questions and important gaps in the field for knowledge producers and/or consumers or for the organizations and systems in which they work.

From the perspective of a knowledge consumer (e.g., an educator or educational leader), curriculum capacity, teacher availability, teacher workload, and education resources can both

support and inhibit the use of knowledge as evidence. Moreover, it is well recognized that knowledge from one setting may not generalize or translate to other educational settings,²¹ and that decision-making structures like curriculum committees, deaneries, and accreditation and licensing bodies can be somewhat removed from classroom or bedside teaching and learning. Decision-makers' attention to institutional and political drivers may limit the use of research evidence as part of their deliberations.²² Take for example, the continued prioritization of student evaluations of teaching, which are known to be flawed and contentious.^{51,52} These evaluations are often used to drive educational practice in a direction favored by students, despite evidence that points to less popular directions that are more appropriate.⁵³ Concerns, such as evidence-user confirmation (i.e., seeking, interpreting, and favoring knowledge that confirms or supports one's prior beliefs or values)⁵⁴ and availability bias (relying on the knowledge that most readily comes to mind), may also be hard to address, particularly when teachers (and learners) believe that what works in education is obvious, common sense, and/or intuitive, making empirical study unnecessary.^{55,56}

The problems of evidence not being valuable or relevant or not mapping well to the complexities of educational governance are not problems to be solved solely by knowledge producers or consumers per se, rather they are part of the bounded rationalities (structural constraints in decision-making) of HPE.⁵⁷ Indeed, it could be argued that both knowledge production and its use as evidence should reflect these realities and idiosyncrasies rather than wishing they could be resolved or were other than what they actually are.

When considering problems at the interface of research and practice, HPE may benefit from better management of the expectations of different stakeholders. For example, knowledge producers might provide a guide to the required timelines and resources for implementing recommended changes so knowledge consumers can plan accordingly. Perhaps, where there are implicit expectations that all evidence generated by HPE scholarship should be about utility and uptake, these expectations could be challenged or at the very least examined more closely. We argue that these expectations could benefit from some nuance and epistemic humility. Not all HPE research is practitioner-oriented, and different kinds of knowledge produced by different kinds of research can serve many other purposes as we described earlier.

Importantly, we cannot expect that individual studies will fundamentally change practice everywhere (or even anywhere). Interpretation, contextualization, and translation are almost always needed.⁵⁸ Indeed, contextual differences between sites, people, interventions, etc., mean that no research finding perfectly generalizes or transfers as evidence from one context to another without consideration of local implementation variables, cultural traditions, and available resources.⁵⁹ Instead, it may be the conceptual artifacts, principles, or philosophies behind those research findings that are most usefully generalizable or translatable to other contexts. Yet, the underlying conceptual artifacts, principles, and philosophies are often poorly articulated by knowledge producers or consumers, likely limiting the ease of uptake or implementation of a variety of educational innovations.^{60–63}

In summary, the problem with evidence may be that too much is expected of HPE knowledge producers and consumers. Can HPE be reconfigured to be more productive, nimble, and dynamic in its approach to evidence-informed practice? Maybe the solutions to HPE problems need to be based on an understanding that they are “the best we can do under the given circumstances” rather than driving for some unattainable and perhaps undesirable ideal. Indeed, what if evidence-informed practice was instead considered to be a systemwide issue in HPE? For example, rather than the current disconnect between the production of scientific knowledge and its uptake, HPE might seek to systematically track the use of such knowledge as evidence. Downloading an article might include a quick survey asking how it will be used as evidence or there might be an automated follow-up alert sometime later asking whether and how the article had been used as evidence (i.e., to inform practice or policy). Another approach could be to systematically conduct reviews of educational innovations and changes (at various levels) to audit what knowledge was used as evidence in what contexts and in what ways. While it might be argued that most instances of translation of knowledge to evidence to practice will not be trackable, if tracking is never attempted, then the problems outlined here will likely continue.

Loop 3: Are the right questions being asked about evidence in HPE?

In this third and final loop, we explore alternative perspectives in thinking about problems with evidence in HPE by asking, “If specific problems with evidence are addressed, then can the broader problems of evidence-informed HPE be subsequently resolved?” After all, one of the biggest limiting factors in HPE is social economics, which focuses on optimizing learning with

limited resources.⁵⁷ Therefore, educational decisions may be based more on compromises than on ideals or evidence. Indeed, given the many constraints HPE programs face, it should not be surprising that evidence often does not play a large role in HPE decision-making.²² After all, the consequences for decision-makers ignoring the recommendations of an academic article are typically much less punitive than for not following legal, institutional, or ethical policies. But what if that changed? What if knowledge producers truly valued the implementation of the knowledge they generate, and knowledge consumers were as accountable for their use of knowledge as evidence as they are for their adherence to policies and procedures?

In loop 2, we argued for more meaningful and realistic connections between evidence (e.g., thinking about what is known and how that supports a claim) and practice. In pursuing these connections, we return to the question of “What is HPE evidence?” and “What purpose(s) is it intended to serve?” The HPE community talks about ‘producing’ evidence, but this could be misleading. As we argued in the introduction, knowledge only becomes evidence when it is used to support or counter a claim, typically involving a decision. Thus, a well-evidenced argument must consider how knowledge can be used as evidence; that is, it must consider the strength of the original knowledge claims and how these knowledge claims can be used to support the argument. One consequence of understanding knowledge used as evidence in terms of utility rather than provenance is that there is no right evidence (whether based on single or multiple sources) independent of an argument. Rather, the quality or “rightness” of knowledge used as evidence is reflected in the way it is used as part of an argument or case,

the appropriateness of the evidentiary inferences drawn from the knowledge, and the ability of those carrying out this work to actually do so.

It is very likely that different knowledge consumers will perceive different utility in different aspects of the same research article. Moreover, utility is not only a matter of applying the right knowledge as evidence. It can also involve the use of knowledge to support claims or positions that it does not really apply to, and it can involve misinterpreting knowledge or using flawed or biased knowledge as evidence. We are not saying this is ideal, only that knowledge consumers use what is useful and that such uses may not be anticipated or acceptable to the producers of this knowledge. Understandably, authors who are advancing knowledge claims of various kinds cannot anticipate how their work will be taken up as evidence nor how their work will be used. However, they can be confident that whatever aspects of their work are used, it will be because those aspects were useful in particular contexts, not because the knowledge was intrinsically right or good. That knowledge can sometimes be misinterpreted or misappropriated when used as evidence suggests that knowledge producers need to be more vigilant about how the knowledge they produce is interpreted, used, or misused, particularly when there are ethical ramifications. Do knowledge consumers also need to be more mindful of their potential to favor knowledge as evidence if it aligns with their beliefs? We argue that mindfulness is needed as a lack of awareness risks concealing blind spots and producing undesirable or even discriminatory outcomes.

The relationship between knowledge production and its use as evidence is, therefore, a relatively decoupled one. This decoupling, in turn, challenges the widespread assumption that education scholars produce and disseminate practice-ready evidence. Our argument is that they produce and disseminate knowledge that might be used as evidence. There is no rule (either tacit or explicit) that knowledge must or should become evidence.⁵ The connections between knowledge (evidence providence) and application (evidence utility) might be strengthened in situations where research is commissioned so that it can be used as evidence (such as in program evaluation) or if the research is a priori designed to target the value propositions of decision-makers. In these cases, we might reasonably expect to see researchers targeting specific kinds of decision-maker interests and values, which, in turn, emphasizes the relational and situated nature of knowledge production and its potential for use as evidence. Another strategy would be for knowledge producers to make better use of alternative forms of dissemination (i.e., not limiting knowledge-sharing to peer-reviewed publications). Sometimes, however, this is undertaken by intermediaries rather than the original authors, suggesting there might be a useful role to be played by knowledge brokers who engage in translating knowledge to evidence in ways that support its uptake.

In summary, researchers who do seek to influence decision-making should put thought and effort into making their research relevant and accessible to decision-makers. However, this is not the whole picture as there are other factors to consider. For instance, when there are misalignments between the speed of research to publication, which tends to be relatively slow, and the speed of decision-making, which tends to be much faster.⁶⁴ Rather than accelerating

the former or slowing down the latter, scholarly inquiry should be considered both as an end in and of itself and as a way of advancing broader conversations about HPE, rather than contributing to specific decision-making processes. As a field, HPE can also be more attentive to the mobilization of research knowledge as evidence through partnering with decision-makers and other knowledge consumers, rather than embracing a “knowledge for knowledge’s sake” stance.

Discussion

Deliberations on problems with evidence are not new to health care or to HPE. It is, however, timely to renew the conversation in HPE, not least because of broader societal debates regarding what counts as knowledge and evidence. Using the analytical framework of single-, double-, and triple-loop learning, we have explored different perspectives on the utility of evidence in HPE, whether there are problems with either or both, what can be done about those problems, and whether the responsibility for the production, dissemination, and uptake of evidence lies with knowledge producers, knowledge consumers, or both. For example, an institution may bemoan the utility of research knowledge in its own educational programming even as it perpetuates a culture of “publish or perish” that encourages volume and velocity in its faculty members’ publication practices.

Our first recommendation is to call for systemic reflection about the nature of knowledge and evidence in HPE. For example, noting that the issue of utility was often seen as contentious in preparing this article, we are not arguing that we should only consider utility in appraising

scientific knowledge, only that utility is an inescapable part of the translation from knowledge to evidence. Indeed, we need to be clear, whether we like it or not, that translation is not bounded by the quality of the knowledge it draws on, as it may engage what might be considered weak or ambiguous knowledge to make disproportionately strong evidential claims. Thus, we ask knowledge producers and consumers to be more attentive to this translation and its transparency and accountability. After all, researchers are careful to track the translation from data to findings within a study, so why not hold the translation from findings to evidence to a similar standard?

Our second recommendation is for systemic tracking and audit of how research knowledge is used as evidence. After all, although our focus has mostly been on the products of formal scientific research and scholarship, we acknowledge there are many other sources of knowledge that may be viewed or used as evidence. And as HPE continues to consider the provenance and utility of evidence, sources of knowledge will increasingly enter the discussion. We also note that our discussion in loop 3 ended at a different place than it did in loops 1 and 2, such that we can contrast the argument that evidence can or should change practice (an understanding of classic translation, which is the bulk of the earlier argument) with the argument that evidence can or should change the way a problem is thought about or discussed. It may be that knowledge producers and consumers can effect change in policy or practice more substantially through changing, adapting, or shifting conversations in HPE than through recommending specific changes in practice.

A third recommendation is for a systemic culture change in how HPE produces and consumes knowledge. At present, it seems the field has a somewhat polarized view of knowledge producers and consumers. Positioning oneself at one or the other of these poles may mask the reality that many members of the community move fluidly between these positions and may undermine opportunities for reflexive dialogue and integration of knowledge producer and consumer goals within programs, institutions, and the field at large. The consequences of this polarized culture are evident, and failure to address this polarization may stifle shared goals of educational progress and improvement and heighten frustration that knowledge lies fallow or is produced but unusable. In suggesting this, we do want to reemphasize the point that research does not always have to serve practical purposes or address the problems of a particular program or institution. That would likely limit innovative and creative ideas and possibilities for the field. We also acknowledge that other systemic change may either facilitate or inhibit (or both) such efforts, such as using generative artificial intelligence to track and report on knowledge-to-evidence-to-practice translations.

Finally, terminology can be misleading. Can the field of health professions education really be reduced to no more than knowledge producers and knowledge consumers? Clearly things are more complicated and intertwined than that. Instead, health professions educators should be asking, “How does knowledge activated as evidence add to or change conversations in HPE?” and “How do HPE institutions shape knowledge production and its use as evidence, and for what purpose?” We have not answered any of these questions definitively, but we hope we

have, by asking compelling questions, encouraged others to think about these issues more deeply.

ACCEPTED

References

1. Asimov I. *The Roving Mind*. Prometheus Books; 1997.
2. Giacomini M. Theory-based medicine and the role of evidence: why the emperor needs new clothes, again. *Perspect Biol Med*. 2009;52(2):234-251.
3. Guyatt G, Cairns J, Churchill D, et al. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420-2425.
4. Mercuri M, Upshur REG. Evidence-based medicine and public health. In: Venkatapuram S, Broadbent A, eds. *The Routledge Handbook of Philosophy of Public Health*. Routledge; 2022.
5. Albert M, Hodges B, Regehr G. Research in medical education: balancing service and science. *Adv Health Sci Educ Theory Pract*. 2007;12(1):103-115.
6. Weiner SJ. From research evidence to context: the challenge of individualising care. *BMJ Evidence-Based Med*. 2004;9(5):132-133.
7. van der Vleuten CP, Dolmans DH, Scherpbier AJ. The need for evidence in education. *Med Teach*. 2000;22(3):246-250.
8. Dauphinee WD, Wood-Dauphinee S. The need for evidence in medical education: the development of best evidence medical education as an opportunity to inform, guide, and sustain medical education research. *Acad Med*. 2004;79(10):925-930.
9. Dornan T, Littlewood S, Margolis SA, Ypinazar V, Scherpbier A, Spencer J. Identification of best evidence in medical education. Case study. *Med Teach*. 2007;29(4):e72-e75.

10. Thistlethwaite J, Davies H, Dornan T, Greenhalgh T, Hammick M, Scalese R. What is evidence? reflections on the AMEE symposium, Vienna, August 2011. *Med Teach*. 2012;34(6):454-457.
11. Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence.' *Med Educ*. 2012;46(1):97-106.
12. Thomas A, Gruppen LD, van der Vleuten C, Chilingaryan G, Amari F, Steinert Y. Use of evidence in health professions education: attitudes, practices, barriers and supports. *Med Teach*. 2019;41(9):1012-1022.
13. Tractenberg RE, Gordon M. Supporting evidence-informed teaching in biomedical and health professions education through knowledge translation: an interdisciplinary literature review. *Teach Learn Med*. 2017;29(3):268-279.
14. Maggio LA, Costello JA, Ninkov AB, Frank JR, Artino AR Jr. The voices of medical education scholarship: describing the published landscape. *Med Educ*. 2023;57(3):280-289.
15. Ioannidis JPA. Educating educators on research on research. *Perspect Med Educ*. 2022;11(3):137-138.
16. Leif M, Semarad N, Ganesan V, et al. The quality of evidence in preclinical medical education literature: a systematic review. *Adv Med Educ Pract*. 2019;10:925-933.
17. Cook DA, Beckman TJ. Reflections on experimental research in medical education. *Adv Health Sci Educ Theory Pract*. 2010;15(3):455-464.
18. Yardley S. Lost in translation: why medical education research must embrace "real-world" complexities. *Med Educ*. 2014;48(3):225-227.

19. Maggio LA, Thomas A, Chen HC, et al. Examining the readiness of best evidence in medical education guides for integration into educational practice: a meta-synthesis. *Perspect Med Educ*. 2018;7:292-301.
20. van Enk A, Regehr G. HPE as a field: implications for the production of compelling knowledge. *Teach Learn Med*. 2018;30(3):337-344.
21. Onyura B, Légaré F, Baker L, et al. Affordances of knowledge translation in medical education: a qualitative exploration of empirical knowledge use among medical educators. *Acad Med*. 2015;90(4):518-524.
22. Thomas A, Ellaway R. Rethinking implementation science for health professional education: a manifesto for change. *Perspect Med Educ*. 2021;10(6):362-368.
23. Harden R, Grant J, Buckley G, Hart IR. Best evidence medical education. *Adv Health Sci Educ Theory Pract*. 2000;5(1):71-90.
24. van der Vleuten CPM, Driessen EW. What would happen to education if we take education evidence seriously? *Perspect Med Educ*. 2014;3(3):222-232.
25. Argyris C. Double-loop learning, teaching, and research. *Acad Manage Learn Educ*. 2002;1(2):206-218.
26. Tosey P, Visser M, Saunders MNK. The origins and conceptualizations of 'triple-loop' learning: a critical review. *Manage Learn*. 2012;43:291-307.
27. Marinopoulos SS, Dorman T, Ratanawongsa N, et al. Effectiveness of continuing medical education. *Evid Rep Technol Assess (Full Rep)*. 2007;149:1-69.

28. Thomas A, Bussi res A. Leveraging knowledge translation and implementation science in the pursuit of evidence informed health professions education. *Adv Health Sci Educ Theory Pract.* 2021;26(3):1157-1171.
29. van der Vleuten CPM. Medical education research: a vibrant community of research and education practice. *Med Educ.* 2014;48(8):761-767.
30. Ratnani I, Fatima S, Abid M, Surani Z, Surani S. Evidence-based medicine: history, review, criticisms, and pitfalls. *Cureus.* 2023;15(2):e35266. doi:10.7759/cureus.35266
31. Doja A, Eady K, Warren A, Wiesenfeld L, Writer H. Utilization of evidence-based tools and medical education literature by Canadian postgraduate program directors in the teaching and assessment of the CanMEDS roles. *Can Med Educ J.* 2021;12(2):e57-e62.
32. Doja A, Lavin Venegas C, Cowley L, Wiesenfeld L, Writer H, Clarkin C. Barriers and facilitators to program directors' use of the medical education literature: a qualitative study. *BMC Med Educ.* 2022;22(1):45.
33. Maggio LA, Larsen K, Thomas A, Costello JA, Artino AR Jr. Scoping reviews in medical education: a scoping review. *Med Educ.* 2021;55(6):689-700.
34. Ellaway RH. Journal standards. *Adv Health Sci Educ.* 2022;27(1):1-5.
35. Mano-Negrin R, Mittman B. Theorising the social within physician decision making. *J Manage Med.* 2001;15(4):259-266.
36. Lucey CR, Hauer KE, Boatright D, Fernandez A. Medical education's wicked problem: achieving equity in assessment for medical learners. *Acad Med.* 2020;95(12 Suppl):S98-S108.

37. Lor P. 2022. Scholarly publishing and peer review in the global south: the role of the reviewer. *JLIS.It.* 2022;14(1):10-29.
38. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. *Acad Med.* 2015;90(8):1067-1076.
39. Horsley T, Steinert Y, Leslie K, Oswald A, Friesen F, Ellaway RH. The use of BEME reviews in the medical education literature. *Med Teach.* 2020;42(10):1171-1178.
40. Chamberland M, Setrakian J, Bergeron L, Varpio L, St-Onge C, Thomas A. Harnessing a knowledge translation framework to implement an undergraduate medical education intervention: a longitudinal study. *Perspect Med Educ.* 2022;11(6):333-340.
41. UK Research and Innovation. How Research England Supports Research Excellence. Accessed March 18, 2024. <https://www.ukri.org/who-we-are/research-england/research-excellence/ref-impact/>
42. Gruppen LD. Improving medical education research. *Teach Learn Med.* 2007;19(4):331-335.
43. Artino AR Jr, Driessen EW, Maggio LA. Ethical shades of gray: international frequency of scientific misconduct and questionable research practices in health professions education. *Acad Med.* 2019;94(1):76-84.
44. Gibbs T, Durning S, Van Der Vleuten C. Theories in medical education: Towards creating a union between educational practice and research traditions. *Med Teach.* 2011;33(3):183-187.

45. Biesta GJJ, van Braak M. Beyond the medical model: thinking differently about medical education and medical education research. *Teach Learn Med.* 2020;32(4):449-456.
46. Hochbaum GM, Sorenson JR, Lorig K. Theory in health education practice. *Health Educ Q.* 1992;19(3):295-313.
47. Ellaway R, Tolsgaard M, Martimianakis MA. What divides us and what unites us? *Adv Health Sci Educ.* 2020;25(5):1019-1023.
48. Paton M, Naidu T, Wyatt TR, et al. Dismantling the master's house: new ways of knowing for equity and social justice in health professions education. *Adv Health Sci Educ Theory Pract.* 2020;25(5):1107-1126.
49. Hawkins RE, Welcher CM, Holmboe ES, et al. Implementation of competency-based medical education: are we addressing the concerns and challenges? *Med Educ.* 2015;49(11):1086-1102.
50. Wartman SA, O'Sullivan PS. The case for a national center for health professions education research. *Acad Med.* 1989; 64(6):295-299.
51. Spooren P, Brockx B, Mortelmans D. On the validity of student evaluation of teaching: the state of the art. *Rev Educ Res.* 2013;83(4):598-642.
52. Ginsburg S, Stroud L. Necessary but insufficient and possibly counterproductive: the complex problem of teaching evaluations. *Acad Med.* 2023;98(3):300-303.
53. Deslauriers L, McCarty LS, Miller K, Kestin G. Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proc Natl Acad Sci.* 2019;116(39):19251-19257.
54. Plous S. *The Psychology of Judgment and Decision Making.* McGraw-Hill: 1993.

55. Gage NL. The obviousness of social and educational research results. *Educ Res.* 1991;20:10-16.
56. Yates GCR. "How obvious": personal reflections on the database of educational psychology and effective teaching research. *Educ Psych.* 2005;25(6):681-700.
57. Cleland JA, Foo J, Ilic D, Maloney S, You Y. "You can't always get what you want...": economic thinking, constrained optimization and health professions education. *Adv Health Sci Educ Theory Pract.* 2020;25(5):1163-1175.
58. Bates J, Ellaway RH. Mapping the dark matter of context: a conceptual scoping review. *Med Educ.* 2016;50(8):807-816.
59. Thomas A, Rochette A, George C, et al. The definitions and conceptualizations of the practice context in the health professions: a scoping review. *J Contin Educ Health Prof.* 2023;43(4 Suppl):S18-S29.
60. Varpio L, Bell R, Hollingworth G, et al. Is transferring an educational innovation actually a process of transformation? *Adv Health Sci Educ Theory Pract.* 2012;17(3):357-367.
61. Horsley T, Regehr G. When are two interventions the same? implications for reporting guidelines in education. *Med Educ.* 2018;52(2):139-147.
62. Norman G. RCT = results confounded and trivial: the perils of grand educational experiments. *Med Educ.* 2003;37(7):582-584.
63. Sherbino J, Regehr G, Dore K, Ginsburg S. Tensions in describing competency-based medical education: a study of Canadian key opinion leaders. *Adv Health Sci Educ Theory Pract.* 2021;26(4):1277-1289.

64. Maggio LA, Bynum WE IV, Schreiber-Gregory DN, Durning SJ, Artino AR Jr. When will I get my paper back? a replication study of publication timelines for health professions education research. *Perspect Med Educ.* 2020;9(3):139-146.

ACCEPTED