[Health Professions Education](https://hpe.researchcommons.org/journal)

Volume 4 | Issue 2 | Article 1

2018-05-15

The Art of Acknowledging that We Know Nearly Nothing

Jimmie Leppink
School of Health Professions Education, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

Recommended Citation

Leppink, Jimmie (2018) "The Art of Acknowledging that We Know Nearly Nothing," *Health Professions Education*: Vol. 4: Iss. 2, Article 1.
DOI: 10.1016/j.hpe.2018.03.004
Available at: https://hpe.researchcommons.org/journal/vol4/iss2/1

Available online at www.sciencedirect.com

ScienceDirect

Health Professions Education 4 (2018) 67–69

www.elsevier.com/locate/hpe

The Art of Acknowledging that We Know Nearly Nothing

Jimmie Leppink

School of Health Professions Education, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

Received 1 March 2018; received in revised form 14 March 2018; accepted 16 March 2018
Available online 17 March 2018

Introduction

It is early March of 2018 – Winter is supposed to be coming to an end and so is a book I have been writing together with two fellow scholars in the field of education. Writing a book, your mind goes back to conference discourse, visiting scholar trips to institutes in other countries, conversations over a good beer, literature you read, things you learned from your teachers, intriguing questions from some of your students, questions that kept you awake at night and remain unanswered, and so much more. The reiterating melody or *Leitmotiv* in all these activities, related or unrelated to the content of the book, is: *scio me nescire* – I know that I know nothing. Any field has its issues and probably there are researchers in medicine, climate studies and other fields who share this *Leitmotiv* with regard to their field, but since I defended a PhD thesis on statistics education in June 2012 and have been working in educational science since then, I will focus on this field. Moreover, since I cannot discuss an entire field of education in one article, I will narrow the focus a bit: A student approaches you with the question how to help medical students develop probability calculus problem-solving skills.

It depends on conditions we do not know (well) yet

Any answer to the question just shared will greatly depend on the respondent's philosophical perspective on learning, the educational research literature one reads, teaching experience, research one is involved in, conversations with peers, and more. One group may prescribe a recipe of fading instructional guidance (i.e., *scaffolding*) through worked examples, completion tasks and ultimately autonomous problem solving. Another group may point at studies that appear to contradict that scaffolding strategy: initial failure on autonomous problem solving *before* worked examples or other direct instruction may sometimes stimulate learning more than starting with direct instruction. A third group may argue that high-frequency low-stakes performance assessments with clear feedback constitutes the best approach. A fourth group may point at studies that have demonstrated benefits of specific forms of group learning over individual learning and will therefore recommend (one of) those forms of group learning. And this non-exhaustive list of possible responses goes on and on. Who is right? It depends on a variety of conditions, some of which we may have started to identify some of which we hardly know anything about.

Learning: a longitudinal phenomenon that is predominantly studied cross-sectionally

One of the reasons why a question like how to help medical students develop probability calculus problem-solving skills may be much more difficult to address than it seems at first is the way learning is commonly studied. Whether we define learning in abstract terms such as *schema development* or in more directly observable terms such as *behavioral change*, learning

E-mail address: jimmie.leppink@maastrichtuniversity.nl

Peer review under responsibility of AMEEMR: the Association for Medical Education in the Eastern Mediterranean Region

https://doi.org/10.1016/j.hpe.2018.03.004

is typically not something that happens in a moment and never fades. Rather, learning is a *longitudinal* phenomenon. Nevertheless, only relatively few studies on learning are longitudinal (e.g., several months) or include repeated measurements in a limited time interval (e.g., posttest immediately after learning and delayed posttest one week later). The question on the development of probability calculus problem-solving skills among medical students is not the same as 'which approach to probability calculus problem-solving skills development results in the best posttest performance immediately after learning?' We are usually more interested in *sustainable* effects of instructional formats on learning, yet we rarely study the sustainability of potential positive effects established initially; as far as I am concerned, until proven otherwise it may well be that an initial positive effect of a treatment fades or even reverses with time. Whether our recipe is fading guidance, letting students struggle first, high-frequency low-stakes performance assessments with clear feedback, a form of group learning, or some combination of the aforementioned options, how confident can we be in our recipe if few if any studies have attempted to investigate the long-term effects of that recipe?

Assessment likely influences learning

Some readers may wonder if I think that all research in education should be longitudinal. My answer to that is 'no'. Whether we ask students questions on the perceived difficulty of a problem or on the effort they invested in a problem or we let them solve probability calculus problems to assess what they learned during treatment, a question recognized by many researchers in education is: can we actually measure or otherwise (e.g., qualitatively) assess learning processes or learning outcomes without influencing these very processes or outcomes? I am inclined towards a '*no*' answer and see an *assessment and learning paradox*: questions on the sustainability of effects of instructional formats may require longitudinal studies, but through repeated measurement or assessment otherwise we may influence that same sustainability of effects in an unknown way, even if we do not explicitly provide feedback to the students who are measured or assessed. To possibly account for such unintended effects, we may need fairly complex experimental designs the feasibility of which is questionable for many reasons.

Back to the Leitmotiv

The goal of this thought piece is not to discard any of the hard labor done by educational researchers across the world; it is just to remind us of how a question that may at first appear simple may quickly turn out complex. I am not advocating to no longer have any confidence in the research we have been doing; we should just bear in mind that interpretations like 'this works (for this type of learner)' and 'that does not work (for this type of learner)' must be made with appropriate caution and may always remain debatable. In the words attributed to the statistician George Box: "*Essentially, all models are wrong, but some are useful*" ([1] p. 424). Models are always a reduction of reality and usefulness is, at least to a large degree, perspective- and condition-dependent. I am using an example from education because this happens to be the main field in which I have operated over the past years. I could have also taken one of the other domains in which I have a background, for instance statistics. I once used to think that every researcher should know 'basics' such as regression and factor analysis. I occasionally laughed about what to me appeared 'silly' mistakes such as calculating Pearson's linear correlation coefficient (r) based on three ($n = 3$) observations or interpreting a nearly 0.5 standard deviations of difference between treatment and control condition in a small-sample experiment as 'no difference' because '$p > 0.05$ hence the null hypothesis is confirmed.' A sample of $n = 3$ is way too small for a meaningful use of r, and the reasoning that '$p < 0.05$, hence the null hypothesis is confirmed' is a *non-sequitur* argument: if a number is based on a particular assumption, it cannot provide evidence in favor of that very assumption. Since p is a probability under the condition that the null hypothesis is true, it cannot provide evidence in favor of the null hypothesis. However, what may seem silly for one may be difficult for the other. For instance, having no medical background, me trying to engage in clinical reasoning would likely end in a straight disaster. Moreover, like other areas in science, education and statistics are complex domains that are in constant movement and development. The more time I spend in these domains, the more I realize how little I know and how much more time I would need to keep track of all the developments in these domains that capture my interest. Whether someone approaches me with a question like how to help medical students develop probability calculus problem-solving skills or how to analyze a particular set of data, I quickly find myself facing a variety of routes that each respond to the question in a somewhat different way.

To conclude

In science, every question and every answer to a question may lead to several other questions. All the questions that are generated in this process guide our research and may lead to some knowledge but can simultaneously make us realize that what we know is essentially just a spark in the universe. Science is not primarily about knowing a lot; as I see it, it is the art of acknowledging how little we know. Through confirmation bias and tunnel vision, attempts to make progress based on assumptions that – based on previous research – we know a lot about *X*, *Y*, and *Z*, may well fail. However, when we take George Box's "*all models are wrong, but some are useful*" as starting point, we realize that we answer to the best of our (limited) knowledge but may well be wrong. This may provide a good remedy against false confidence either on the part of scientists or on the part of consumers. Neither scientists nor consumers should equate empirical evidence in favor of a treatment established in a particular study with 'this treatment works (for everyone, no need to consider any conditions)'. Moreover, taking Box's words as starting point may also help us to create and preserve a safe learning and working environment in which both experienced and less experienced actors may make mistakes and, through dialogue, learn from these mistakes.

Ethical approval

Not applicable.

Funding

None.

Other disclosures

No conflicts of interests.

Reference

1. Box GEP, Draper NR. *Empirical Model-building and Response Surfaces*. New York: Wiley; 1987.