

The future of practice-based research in educational technology: Small steps to improve generalisability of research

Sakinah S. J. Alhadad

Griffith University

Implicit in the discourse of evidence-based practice are two fundamental concerns. One is the generalisability of research evidence where issues of external validity are integral to translation, relevance, and application in complex and multifaceted higher educational contexts. The other relates to practice-based evidence, where issues of internal validity impact on the design, interpretation, and dissemination of research. While practice-based research has an advantage in terms of high external validity, threats to internal validity can cause significant issues in terms of the subsequent inference, translation, and generalisability of findings. In educational technology, evaluation and research of e-learning in higher education is conducted by both practitioners and academics, each contributing different pieces of the puzzle towards a better understanding of the learning processes in complex real world settings. In this paper, I propose small, practical steps towards improving the generalisability of practice-based research.

Keywords: Practice-based research, evaluation research, research methods, validity, measurement, generalisability

The future of practice-based research: An evolving journey

This paper is a product of reflections precipitated by my relatively recent move into an area of work that is at the nexus of research and practice. Due to my background as an educator and researcher in the psychological science domain, the challenges I currently face at this nexus trigger a self-motivated need to enhance the voice and value of practice-based research, specifically in the applicability and dissemination of education research in general (both basic and applied). As I begin the journey into an authentic understanding of fit-for-purpose practitioner research, these are my preliminary thoughts towards enhancing the generalisability of practice-based research. The principles of the suggestions made here are design- and measurement-agnostic, in the hope that practice-based researchers will be able to apply these to their contexts, whether they are involved in design research, focus group evaluation, or learning analytics research. The purpose of this paper is not to simply share my exploratory intellectual journey in this space, but to engage peers (both juniors and seniors) and leaders in an effort to contribute towards the enhancement of educational research in the broader community of others also living in and around the research-practice nexus.

Practice-based research as educational technology research

Evidence-based practice operates at two fundamental levels: the first is to *use* existing evidence and apply it to the practice; the second is to *establish* evidence where gaps exist in the current evidence base, or where existing evidence may be questionable, weak, or uncertain (Davies, 1999). In this paper, I focus on the latter. Practice-based research in educational technology (analogous to 'evaluation research' see Philips, Kennedy, & McNaught, 2011), provides macro- and micro- level views of teaching and learning-related phenomenon in dynamic and complex settings, which can in-turn provide a very rich source of directions for hypotheses for more controlled empirical research. The synergistic effect of both controlled empirical research and practice-based research increases the likelihood of studying the educational phenomenon more authentically. In this paper, I make some practical suggestions to optimise the balance of internal and external validity with the aim of enhancing the generalisability of this research (replicability/applicability in practice, in different contexts and populations of research findings in this domain). Other pertinent issues related to practice-based research are out of the scope of this paper and are covered elsewhere (for comprehensive guidelines and discourse on evaluation research at each stage of the e-learning

lifecycle, see Phillips, et. al, 2011; for critical discourse on the role of theory in evaluation research, see Hannon & Al-Mahmood, 2014, and Phillips, Kennedy, & McNaught, 2012).

What do issues of internal and external validity mean for practice-based research?

Campbell (1957) introduced the concept of internal and external validity as a means to evaluate the value of experimental designs in social settings. Internal and external validity in research design are often conflicting ideals. Optimising a research design for high internal validity incurs trade-offs for external validity, and vice versa. For any given research design, resources are limited, and this makes it very difficult to identify and measure all variables that may be influencing the observed effect in any given research scenario. The question is, how do we make design decisions to optimise internal and external validity, and on what do we place emphasis when considering the purpose of our research?

Broadly, external validity addresses the question of whether a particular finding is generalisable across a variety of contexts, settings, persons, and times. External validity in Campbell's (1957) original definition referred to the generalisability of the studied effect (and of its underlying processes) across different participants, settings, and research methods. This was later distinguished into: (1) ecological validity, or the degree to which the research design replicates the actual occurrence of the scenario/circumstance in naturalistic settings; and (2) relevance of generalisability, or the degree to which the research findings can be generalised across different participant populations, contexts, and other related settings (Brewer, 2000). A research design with high external validity necessarily closely resembles or ideally replicates the authentic experience in the authentic setting. A benefit of having a design with high external validity is that any research finding can be seen to be generalisable to a real context. To give an example, observational research in real-world settings gathers genuine data on observable behaviours, and thus can be argued to be representative of real-world behaviours. Such a research approach has strength in identifying existing naturalistic relationships. However, optimising a research design for maximum external validity limits the degree to which genuine cause-and-effect relationships can be identified, owing to the impossibility of measuring the influence of the large number of variables in the naturalistic setting that may be influencing the target phenomenon.

Internal validity refers to the extent to which we can accurately infer or conclude that the independent variable (or predictor) produced the observed effect on the dependent variable (or criterion) (Campbell, 1957). That is, internal validity is the degree to which we have confidence that a true causal relationship exists. In experimental research (often lab-based), ascertaining whether the observed effect (as measured by the dependent variable), is truly caused by or predicted by the independent variable is relatively less challenging than if it were tested in an applied environment. Determining confidently that the observed effect is solely a function of the one independent variable is dependent on controlling other potential influencing factors (i.e. extraneous variables: variables other than the one being investigated). In the controlled experimental context it is possible, to an extent, to keep the influence of extraneous variables (particularly ones that impact on the relationship systematically, named 'confounds') constant, consequently eliminating any differential influence these variables may have across the various levels of the independent variable. This ability to tightly control variables to optimise internal validity is the strength of lab-based research, but this approach is very difficult to implement in natural settings such as those typical in educational technology research. Further, when intending to optimise experimental control over extraneous variables it is necessary to recognise that variables are not equal – some variables are easier to control than others. Designing experiments that generalise within and across contexts involves a complex interplay of internal and external validity. Below I suggest ways to improve internal of external validity of practice-based research, and to increase the understanding of this psychometric property in research within complex environments.

The way forward: Practical suggestions to enhance generalisability

These strategies are suggested with the aim to increase the translation, utility, and application of the research *outside* of the research context, and to enhance the efficiency of the evaluative design process *within* context.

- 1. Nuance your evaluation research questions to increase the understanding of why or how the causal effect or relationship works in context**

1.1. Consider factoring in control or comparison groups.

In the academic capacity development domain, a frequently identified issue when evaluating academics' practice-based classroom research is that claims of efficacy of intervention or effects studied often exclude an appropriate control condition or comparison group. (e.g., Benassi et al., 2014). Being thorough in identifying drivers of an effect requires appropriate controls in place to enable more accurate inferences or conclusions to be drawn. For example, in order to infer whether learning strategy *A* (e.g., structured reflective practice) was effective, having a control condition *B* such as, in this case, 'unstructured reflective practice', allows the researcher to more accurately infer that the pattern of findings was not merely due of the act of reflective practice, but because strategy *A* was structured for optimal reflection for the task or goal.

Often in practice-based research, it may be either impossible or unethical to have the ideal control group. In this case, a strategy the researcher could adopt is to statistically control the measurable confounds or influencing variables (within reason). By quantifying variables or factors that may have influenced the observed effects, internal validity is enhanced as it now creates new plausible hypotheses. Using the example above, if having a control condition was not logistically possible, one may be able to quantify the amount of engagement with the act of structured reflection, and thus be able to quantify whether the extent of engagement with the reflection task impacts systematically on the learning outcome. Another example might be to measure related psychological constructs as covariates in the model (also see point 1.2 below); whether or not the student is a deep or surface learner may impact on the magnitude of effect observed as a whole group, so quantifying this will enable the researcher to understand more deeply the mechanisms behind successful adoption of this task, and how to further improve the design and application subsequently.

1.2 Use theory as a means to frame research design.

Theories or models used as a research design framework can really enrich practice-based research design and hypotheses (see Figure 1 for example; see also Lizzio, Wilson, & Simons, 2002; Kember, McNaught, Chong, Lam, & Cheng, 2010). Mook (1983) purported that the component of an experiment that increases its generalisability capital is the theoretical process or understanding that accrues from the study. In the Figure 1 example, I suggest ways to use an example of an omnibus model, the 3P model (Biggs, 1989; 1993), to nuance your research question to enhance internal validity and generalisability of the research. The 3P model comprises three main components representing an integrated system of student learning: *Presage*, or pre-existing, (relatively) stable student characteristics that relate to learning, and to the instructional context; *Process*, the underlying factors related to the process of the learning task itself; and *Product*, the learning-related outcomes.

Where research designs begin with the *Presage* and *Product* stages, a way to nuance the research question further is to look at the moderating or mediating effect of *Process* factors. For example, if the product/intervention leads to enhanced academic achievement, does this pattern of results change as a function of whether the students are high or low on self-efficacy? This nuanced question allows for more efficient refinement of either the design or the investigated intervention/product as a result of clearer understanding of why the effect is occurring. For the practitioner, this also provides an actionable strategy for design improvement in the classroom or curriculum design. This iterative cycle of the design and practice-based research process is indicated by the green arrow in Figure 1 (bottom panel).

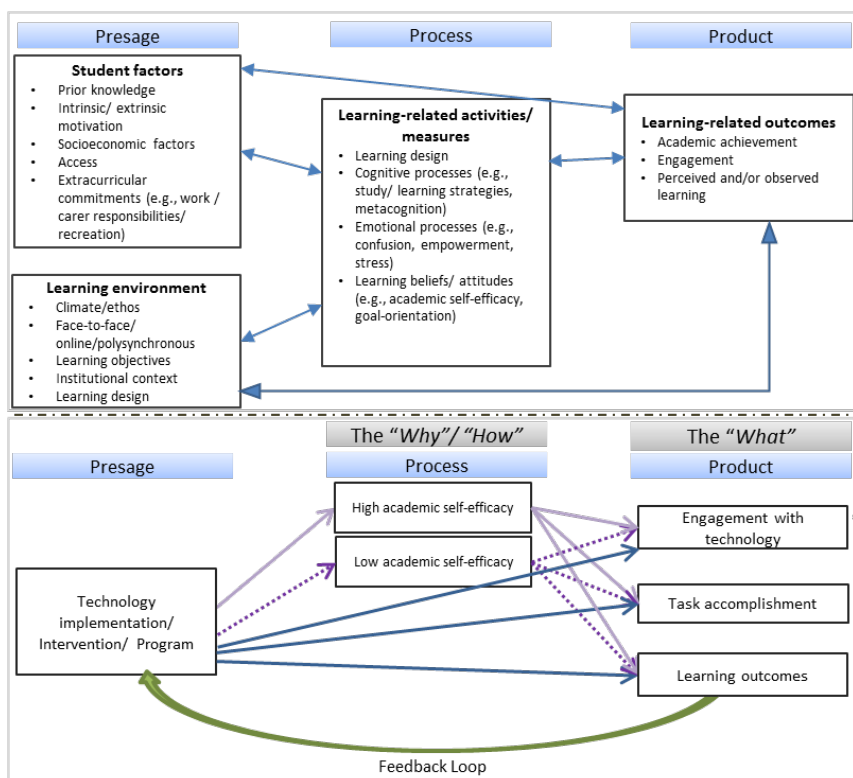


Figure 1. Adapted from Biggs (1989).

Top panel: Ways to enhance the 3P model as a framework for practice-based research. Bottom panel: * “Engagement with technology” here, can be either the Process (in this case, it can also act as a covariate) or Product (outcome), depending on the research question. Each factor can also be measured in multiple ways; carefully assess your environment as to what is available and possible to measure. For example, the uptake of technology may be measured through learning analytics (use/do not use, frequency of use, pattern of use over time). As a process, the critical design thinking may be centred around the questions you *could* ask – what data sources do you have access to; what is pragmatically measurable? For example, the question can change to: does the impact of technology use on learning change as a function of the frequency of use? As such, “engagement with technology” is now a covariate in this model.

2. Consider moving beyond student perceptions: convergent measures

The likelihood of accurately measuring student perceptions as they relate to the intended effect or construct depends on the research question. If, for example, you are interested in assessing the effectiveness of an educational technology in terms of usability, asking students for their perceptions of their own attitudes and beliefs is appropriate, and is likely to be an accurate representation of the true effect (Note: for a good primer resource on survey/questionnaire development, see DeVellis, 2012). However, if you are interested in effectiveness or learning-related outcomes, asking students for their perceptions will give a false impression of the intervention’s or product’s effectiveness (Phillips et al., 2011). Findings pointing to similar conclusions are plentiful in the cognitive science literature – when students are asked to judge their own level of learning during study or in a test, students tend to misjudge their actual learning performance (e.g., Asher & Bjork, 2005; Castel, McCabe, & Roediger, 2007). Including learning-related process measures such as study strategies, engagement in formative assessment as aligned with learning design, or proxies of engagement or effort such as various sources of learning analytics (see Lockyer, Heathcote, & Dawson, 2013) will enable deeper explanations and inference of the learning phenomenon studied in context. In Figure 1 (Bottom panel) I suggest ways to move beyond student perceptions in evaluating the effectiveness of a product/intervention/design by including other dependent measures to be converged with measures of perceptions to increase our understanding of the studied phenomenon. Use of self-report is beneficial here if the aim is to assess attitudes or perceptions such as ease of use, satisfaction with use of technology for learning, and perceived development of skills and learning.

3. Consider the research questions you want answered and how to statistically test them before collecting data.

Practical research is difficult to initiate in the first instance so opportunities to gather data need to be exploited. Consider how you would answer your research questions with statistics, and iteratively evaluate the levels of measurements of your dependent variables to ensure you have optimally designed your measures to answer your questions. Further, where quantifiable, report effect sizes as complementary to quantitative statistics. The size of the magnitude of the findings in the research may be reported if the learning-related dependent variables are measured on a quantitative scale. This allows for standardised comparison of the observed effect in terms of its quantified magnitude across studies (see for Cohen, 1992 for a primer; Cumming & Finch, 2001, or Cumming, 2012 for more depth; and Hattie, 2009 for discussion specific to quantification of learning measures).

4. Communication or dissemination: Report important information on contextual variables

The overarching principle in enhancing generalisability of practice-based research studies is in the communication of the research findings and the details that facilitate generalisability. Be cognisant of, and acknowledge context specificity of the findings. Explicitly address the external and local realities in communication and dissemination of practice-based research. (Green, 2008). Important information on contextual settings such as representativeness of sample, reach, implementation methods, and other pertinent variables would help readers in assessing more accurately the applicability of the study results to their own context (Glasgow et al., 2006).

Conclusion

In this paper, I offer small, practical ways to optimise the balance of internal and external validity to facilitate the design, dissemination, and applicability of practice-based research. The utility of these suggestions are meant to be paradigm-agnostic, however the goodness of fit will often be less than ideal, as various factors in the multifaceted, dynamic, complex environments of practice-based researchers will interact differently. In the interpretation and application of research evidence, these principles are equally important in maintaining an appropriate level of skepticism and in establishing the quality and accuracy of inference to future research and practice, such that one is able to prevent acceptance and replication of poorly tested interventions or research. The strategies recommended in this paper are an effort to contribute towards the enhancement of educational research in the broader community of others also living in and around the research-practice nexus.

References

- Asher, K. & Bjork, R. A. (2005). Illusions of competence in monitoring one's own knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 187-194.
- Benassi, V. A. Tappin, E. M., Overson, C. E., Lee, M. J., O'Brien, E. J. White, B. P., et al. (2014). Applying the science of learning: The cognition toolbox. In V. A. Benassi, C. E., Overson, & C. M. Hakala (Eds.), *Applying the Science of Learning in education: Infusing psychological science into curriculum* (pp194-205).
- Biggs, J.B. (1989). Approaches to the enhancement of tertiary teaching. *Higher Education Research and Development*, 8, 7–25. <https://doi.org/10.1080/0729436890080102>
- Biggs, J. B. (1993). From theory to practice: A cognitive systems approach, *Higher Education Research & Development*, 12(1), 73-85. <https://doi.org/10.1080/0729436930120107>
- Brewer, M. B. (2000). Research design and issues of validity. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 3-16). New York, US: Cambridge University Press.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54 (4), 297-312. <https://doi.org/10.1037/h0040950>
- Castel, A. D., McCabe, D. P., & Roediger, H. L. (2007). Illusions of competence and overestimations of associative memory for identical items: Evidence from judgments of learning and encoding fluency. *Psychonomic Bulletin & Review*, 14, 107-111. <https://doi.org/10.3758/BF03194036>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037/0033-2909.112.1.155>

- Cumming, G. & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, 61, 530-572. <https://doi.org/10.1177/0013164401614002>
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge. <https://doi.org/10.4324/9780203807002>
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47 (2), 108-121. <https://doi.org/10.1111/1467-8527.00106>
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. California: Thousand Oak.
- Glasgow, R.E., Green, L. W., Klesges, L. M., Abrams, D. B., Fisher, E. B., Goldstein, M. G., et. al. (2006). External validity: We need to do more. *Annals of Behavioral Medicine*, 31 (2), 105-108.
- Green, L. W. (2008). Making research relevant: If it is an evidence based-practice, where's the practice based evidence? *Family Practice*, 25 (supp) , i20-i24.
- Hannon, J., & Al-Mahmood, R. (2014). The place of theory in educational technology research. In B. Hegarty, J. McDonald, & S.K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014* (pp. 745-750).
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
- Kember, D., McNaught, C., Chong, F. C. Y., Lam, P., & Cheng, K. F. (2010). Understanding the ways in which design features of educational websites impact upon student learning outcomes in blended learning environments. *Computers & Education*, 55, 1183-1192.
- Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perception of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education*, 27(1), 27-52. <https://doi.org/10.1080/03075070120099359>
- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, 57 (10), 1439-1459.
- Mook (1983). In defense of external validity. *American Psychologist*, 38(4), 379-387.
- Phillips, R. A., McNaught, C. & Kennedy, G. (2011). *Evaluating e-learning: Guiding research and practice*. New York and London: Routledge <https://doi.org/10.4324/9780203813362>
- Phillips, R., Kennedy, G. & McNaught, C. (2012). The role of theory in learning technology evaluation research. *Australasian Journal of Educational Technology*, 28(7), 1103-1118. <http://www.ascilite.org.au/ajet/ajet28/phillips.html>

Alhadad, S.S.J. (2015). The future of practice-based research in educational technology: Small steps to improve generalisability of research. In T. Reiners, B.R. von Kinsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Globally connected, digitally enabled*. Proceedings ascilite 2015 in Perth (pp. 363-368). <https://doi.org/10.14742/apubs.2015.956>

Note: All published papers are refereed, having undergone a double-blind peer-review process.



The author(s) assign a Creative Commons by attribution licence enabling others to distribute, remix, tweak, and build upon their work, even commercially, as long as credit is given to the author(s) for the original creation.