WESLEY IMMS, BENJAMIN CLEVELAND AND KENN FISHER

# 1. PURSUING THAT ELUSIVE EVIDENCE ABOUT WHAT WORKS IN LEARNING ENVIRONMENT DESIGN

It is an exciting time to be involved in education. Every day we witness the pursuit of innovation and creativity in schools, the sophisticated development of personalized learning approaches, the increasing usefulness of ubiquitous technology, and the excitement surrounding the many ways education can contribute to burgeoning 'knowledge economies'. These 'new age' priorities are combining to enable students to increasingly take responsibility for their own learning and are encouraging teachers to become the curators of learning experiences, that range from whole class didactic encounters, through collaborative peer-peer active learning to reflective one-on-one consultations with students – often within a single lesson.

To meet this change, schools are altering their architecture and spatial arrangements, sometimes quite dramatically. Through discussions between designers and teachers about how learning and teaching should be approached in the 21st century, the traditional corridor and classroom layout is increasingly being reconfigured into 'flexible' school designs. Differing in important ways from the 1970s open classroom and 'free-range learning' concept, the best of these spaces can theoretically accommodate a variety of teacher epistemologies and a range of students' preferred learning styles.

A formidable variety of spatial typologies are now emerging. Dovey and Fisher (2014) summarise these into five genres of design (Figure 1), ranging from the traditional egg-crate style, through to large open space configurations. As represented in the diagram, 'openness' increases from left-to-right across the typologies, with the experienced educator often commenting that maximum flexibility occurs in types C and D, where walls, doors, furniture and spaces can be (re)configured to support a wide array of desired learning and teaching practices, activities and behaviours.

Dovey and Fisher (2014) importantly note that while their typologies describe design trends, they can be better understood as *assemblages*, where iterative practices mix with adaptive space configurations to create a hybrid of space and pedagogy unique to each educative occasion.

*W. Imms et al. (Eds.), Evaluating Learning Environments, 3–18.*

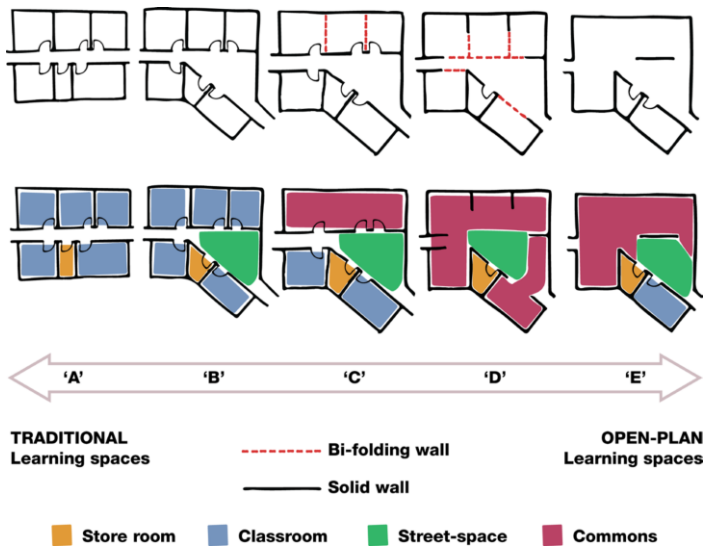


*Figure 1. Dovey and Fisher's learning space typologies (2014), adapted by Soccio & Cleveland, 2015*

Thus, a well-designed open area, such as shown in Figure 2, can simultaneously be a lecture theatre, a setting for small group meetings, a place for visible teaching and learning for staff and students on the periphery (Hattie, 2011), and a learning zone for those transiting between experiences.

*Figure 2. Yarra Valley Grammar School – Science and Mathematics Building, Hayball with Engaging Spaces. Photo copyright Dianna Snape (used with permission)*

Likewise a space can be inhabited by students (Figure 3), a type of colonization where the most important output is the collaborative, democratic synthesis of ideas and knowledge not unlike the ideal described by Dewey (1916/1997) a hundred years ago.

*Figure 3. Coatesville Primary School, Clarke Hopkins Clarke Architects (used with permission)*

Such spatial developments are disruptive interventions that often displace, rather than replace, the teacher. By displacement, we mean that in such socio-spatial settings, the teacher is removed from being the focus of attention with students seeking out the educator they feel owns the specialised knowledge that will best inform their learning issue. While didactic instruction remains a necessary pedagogy, increasingly spaces in schools are being designed and used to allow students to work with a range of knowledgeable others to socially construct knowledge and meaning. As illustrated in Figure 4, students can work in isolation or in small groups. They can use fellow students for peer instruction, undertake independent work in a collegial environment, and utilise the capacity for movement within the space to seek out those who can best inform a learning issue. Rather than removing the teacher, such spaces are designed to elevate them to the role of expert collaborator.

Such scenarios describe an idealized educational world. The figures shown above are staged for the photographs to illustrate how a space is *intended* to be used. And while these photos depict a static situation, frozen in time, they hide the temporal issue which is another facet to be considered in innovative learning environments – when is didactic teaching required, when reflective, when peer-peer-based, when maker-space, when studio/workshop/lab/outdoor learning? Can the spaces adapt to these requirements?

The question remains, when looking at photos such as these, what is the reality? Are the typologies described by Dovey and Fisher (2014) being used in the democratic and differentiated manner described by Dewey (1916/1997) and others? If they are, the logical question is what is their effect? (Hattie, 2008); do they improve teaching and learning standards to the degree where their cost and use is warranted?

*Figure 4. Catholic College Wodonga – Mercy Senior Learning Centre, No. 42 Architects with Engaging Spaces (used with permission)*

As shown in the figures above, what may be termed new generation learning environments (NGLEs) are characterised by polycentric room designs, infused information and communication technologies, flexibility brought about by moveable walls and other agile interior elements, a variety of 'student friendly' furniture, and ready access to resources. They significantly expand our conceptualisation of the learning and teaching space, but bring with them a perplexing problem: what methodologies are 'best fits' for evaluating such environments? The following chapters in *Snapshots* explore the complex issue of how we evaluate the impact of learning environments. In particular they problematize the issue of how we measure the impact of NGLE's on student learning, teacher pedagogy, and associated variables. Indeed Dovey and Fisher (2014) have called these NGLE's *complex adaptive assemblages*, to illustrate the difficulty of evaluating such innovative spaces.

*Snapshots* constitutes new knowledge that will assist teachers to utilise these spaces effectively. Although the re-conceptualising and inhabiting of new school

architecture is moving at an unprecedented pace, teachers' abilities to utilise new spaces are not always matching this growth. An analysis of literature reviews (Cleveland & Fisher, 2014) identifies a general acceptance that many teachers have poor 'environmental competence', thus limited capacity know how to "… understand and effectively use physical instructional space for a pedagogical advantage" (Lackney, 2008). Increasingly, teachers are being challenged to re-think how they teach in order to maximise the instructional use of new learning environments. This raises the notion of teacher *spatial literacy*, which is not to be confused with visual literacy. Spatial literacy, which works in the 3rd dimension and indeed the fourth dimension – the temporal – is also critical to these NGLE's (New London Group, 1996).

Using NGLEs effectively requires a significant evidence base to assist teachers to reconceptualise space as a pedagogic tool (Cleveland, 2011). In addition, new designs are being put into place with scant evidence that the resulting expense and disruption to teaching practice is underpinned by evidence that the designs 'work'. Even a brief review of research from the 'open classroom' era in the 1970s (Imms, 2016) shows that little robust evidence was collected concerning the impact of those spaces on student learning. Without such research, education is doomed to repeat errors of the past. What is required, and quickly, is a robust approach to evaluating the impact of NGLEs. *Snapshots* is a significant move in that direction.

## LINKING EVALUATION THEORIES TO LEARNING ENVIRONMENT CONTEXTS

The *Evaluating 21st Century Learning Environments* (E21LE) Australian Research Council (ARC) Linkage Project was set up to conceptualise, develop and trial innovative approaches to the evaluation of physical learning environments in today's secondary schools. Sponsored by the Australian Federal Government and situated within the Learning Environments Applied Research Network (LEaRN) at the University of Melbourne, the project brought together five industry and educational Partner Organisations to tackle the elusive goal of determining 'what works' in learning environment design. The Partners of the project are indicative of the multi-disciplinary nature of the issue being addressed. An architectural firm (Hayball) and a technology supplier (Keepad Interactive) have formed an alliance with three Australian schools – the Australian Science and Mathematics School in Adelaide, South Australia; the Anglican Church Grammar School in Brisbane, Queensland; and Caulfield Grammar School in Melbourne, Victoria. Led by a team of academics from the University's Melbourne Graduate School of Education and Faculty of Architecture, Building and Planning, and incorporating three PhD candidates (two from Education, one from Architecture) this uniquely multi-disciplinary team is bringing an array of specialist knowledge to this issue.

E21LE is not an evaluation *per se*. Its purpose is to develop a robust evaluation framework that will allow education and design professionals to assess a variety of design and educational variables across a variety of sites for a variety of purposes.

The driving philosophy is that a learning environment evaluation framework must be many things to many people. If one common purpose can be described, it would be to facilitate the collection of rigorous, useable evidence concerning the impact of spatial design on pedagogy and learning. E21LE is conceptualising, developing and trialling this framework through a variety of primary and secondary research methods.

The project is not looking for a single definition or model of evaluation that will direct its practices. Rather, it seeks to map an evaluation 'terrain' that logically links its acts of evaluation to learning space phenomenon (Imms, Cleveland, Mitcheltree, & Fisher, 2015). E21LE's protagonists acknowledge that the research must work within a frame of reference that is meaningful and coherent across the whole education/design landscape. Such an ambiguous foundation is not uncommon. Love (2010) argues that a single approach to evaluation is not often possible and that evaluation is mostly an applied activity that takes place in real-world settings. Likewise, Carden and Alkin's (2012) analysis of evaluation theory recognises that a characteristic of evaluation practice is the making of concessions. Evaluators commonly encounter multiple and competing purposes for evaluation, numerous stakeholders with contradictory needs, limited time and budget for evaluation, and conflicting views about evaluation methodology. A single methodology or method is, under these circumstances, not practical.

The issue of multiple and competing complexities in evaluation is especially pertinent to E21LE. The project is deliberately multi-disciplinary, trying to integrate the sometimes contradictory epistemologies that underpin architecture and education. These disciplines each have their own rich tradition of evaluation, but there are few models that bring these fields together in a consistent and coherent manner. E21LE also spans the infamous research 'paradigm divide' – qualitative vs. quantitative methods – with both seen as being equally useable within the scope of the research questions being posed by this project.

Finally, the research is further complicated by the need to address multiple audiences and supply findings that have applications in architecture, education and public policy. It is clear that the ways in which E21LE defines and then utilises evaluation has *theoretical*, *methodological* and *application*/*audience* perspectives. As such, it is logical to organise this book around an aligned three-part framework: *emerging issues*, *emerging methodologies*, and *emerging knowledge*. This will be described presently.

## A NOTE ON EVALUATION THEORY

Evaluation theories define a project's guiding principles, as compared to its actual practices. These principles define the body of knowledge that '…organises, categorises, describes, predicts, explains, and otherwise aids in understanding and controlling…' the focus of that evaluation (Shadish, Cook, & Leviton, 1991). Theory allows researchers to decide and justify the 'where, when and why' of the application of evaluation methods.

To this degree, the role of theory within evaluation is clear and logical. However, as a relatively new and rapidly developing discipline there exists no common agreement regarding the nature of evaluation's actual structure. For example, theories can be *prescriptive* by stipulating the rules and frameworks that control evaluation; or they can be *descriptive*, providing explanations and statements that generalise possible approaches. Theories can be based within a *realistic* epistemology (Pawson & Tilley, 1997) where it is argued that the agents that dictate outcomes must be the focus of the evaluation; or they can be *theory of change* oriented, where theories concentrate on the actual result and resultant claims of causation (Weiss, 1998). It is possible to have theories *of* evaluation that direct explain and justify an approach, and theories *for* evaluation which address issues of application of the evaluation's aims and outcomes (Astbury, 2012). These elementary distinctions regarding evaluation orientation are discussed later in this paper, where they inform E21LE's positioning of evaluation within its research context.

The question can be asked, 'Is there some logical schema to evaluation theory'? As methodological approaches to evaluation proliferated, Alkin and Christie (2004) sought to give the field some coherence. They developed a 'roots' taxonomy that argued all evaluation methods were logically linked to two core evaluation functions: *accountability & control*, and *social inquiry* (a third was added in 2012: *epistemologies*). They argue that (1) accountability meets a community need to justify investment and ensure future program quality, and (2) social inquiry provides robust methods of systematic evaluation. While the accountability function focuses on *rationales* for evaluation, the social inquiry function focuses on *applications* of evaluation.

Three practices of evaluation have grown from these two foundations: those of *use*, *methods*, and *valuing*. Evaluations with a 'use' practice grew from an accountability foundation and are those focused on the further application of the knowledge gained, as characterised by the theories of Stufflebeam (1983). Evaluations with a 'methods' practice are those that adhere to prescribed research methodologies in order to provide robust findings, characterised by the theories of Campbell (1957). Evaluations with a 'value' practice are those that recognise the role of human judgement in evaluation, characterised by the theories of Scriven (1967). Within this latter category, an 'objectivist' sub-branch limits valuing to being the informed view of the evaluator, while a 'subjectivist' sub-branch recognises the phenomenological nature of evaluation, drawing on the opinions and interpretations of participants. Both the 'method' and 'value' practices had their origins in the social inquiry foundation.

Alkin and Christie (2004) argued that the roots metaphor supported a logical theory 'flow' that applied to all evaluation methods in practice. They argued that such methods could be traced through practices of evaluation (the branches) directly to evaluation's foundational roots. A later paper by Carden and Alkin (2012) added a third foundation of *epistemologies*. They argued the robustness of this taxonomy, stressing that the theoretical 'flow' was not exclusionary, rather that particular

orientations merely reflect the 'primary emphasis' of the evaluator, while accepting methodological and theoretical concessions necessary in any evaluation.

In other words, while theorists can identify specific genres of evaluation, these must be recognised as flexible and adaptive when *in practice*. Practice, though, must remain accountable to the predominant theory that informs it. These two components of evaluation – theory and practice – are interdependent. 'Best practice' in evaluation cannot be developed separate from the theories that inform that practice. As outlined by Chelimsky (2013, p. 91), "each one learns from the other and, in that learning process, both are inspired to stretch, to bend a little, and to grow".

The third element in the evaluation process is *applicability*. While designing an evaluation – and subsequently gathering of data to provide an evaluative analysis – are both critical so, too is how those data are used. Results must serve a purpose and these purposes reflect political functions and social implications. Chelimsky (2013) put it thus:

> … the evaluations we produce, that are based on theory and performed in the real world, are also planned, implemented, analyzed, and reported by people. And those evaluations are open, from beginning to end, to political pressures by policy makers, planners, administrators, special interest groups, subject-area practitioners, participants, and all those who may be affected by the results—or feared results—of the evaluation. (p. 92)

Therefore, with this somewhat ephemeral mixture of theory and practice, evaluation must be approached with caution.

## SITUATING THEORIES OF EVALUATION WITHIN E21LE

What then are the implications of this brief summary for *Snapshots*? In short, it provides a structure for an evaluative framework suitable for the complex phenomenon of innovative learning environments. Thus, these chapters are seen as contributors to the population of that matrix. To explain further, the general purpose of evaluation is to establish *merit*, in its widest societal sense determining anything from the best restaurant or seats at a football game, to informing major government decisions. While approaches to evaluation may vary widely within research, Weiss (1998) claims that all contain five common elements. Evaluation research provides; rigor, through being *systematic*; it focuses on *process* and *outcomes*; it is *comparative*, using established goals against actual outcomes; and it has *purpose* in that it seeks to improve practice. Thus, while evaluation research involves establishing the merit of social initiatives, *its purpose is improvement of future practice*.

Weiss' summary of the purpose of evaluation mirrors the traditionalist approach evident in work done by historical figures such as Cooke, Campbell (in Shaddish et al., 1991), Pawson and Tilley (1997) and others. Through experimental and quasi-experimental evaluations of organizations and programs (see, also Scriven,

1967), evaluation aims to audit existing practices to improve society and generate new theory. Mark, Henry and Julnes (1999) describe this as a representational evaluation purpose. They go further to offer an additional purpose of evaluation, the valuative, where qualities people attribute to phenomena lead them to 'natural' assessments. Mark et al. characterize this as *assisted sensemaking*: the acknowledgement of humans' pre-disposition to make judgments about their existence. The premise here is that the underlying purpose of evaluation is to help, rather than replace, that natural process. A useful way to bring these descriptions together is describing evaluation as serving the purpose of improving future practice by looking back (auditing/appraisal), looking to the future (improvement and prediction/analysis) and looking within (valuing/judgement).

This review of evaluation literature (see, for instance Shufflebeam and Shrinkfield, 2007) would indicate that evaluation research meets the needs of those who wish to *describe* (assess an observable attribute), those who wish to *classify* (assess underlying structures and categories), those who wish to identify *causality* (assess what outcomes can be attributed to a program), and – to add Mark et al. (1999) – those who wish to understand *values* (assess the experiential quality of a program).

As these needs are neither mutually exclusive nor irrevocably linked: the evaluation theories outlined above led the E21LE research team to the concept of an evaluation matrix. Figure 5 represents the project's approach to tailoring evaluative approaches to meet a variety of purposes and needs.

| | **Meeting the *needs* of those who wish…** | | | | | |
|---|---|---|---|---|---|---|
| | | | To describe | To classify | To identify causality | To determine value |
| | (Aim) | (Through) | | | | |
| **The *purpose* of evaluation research is…** | | | | | | |
| | To improve | Formative analysis (judgment) | | | | |
| | To audit | Summative analysis (appraisal) | | | | |
| | To generate theory | Predictive analysis (analysis) | | | | |

*Figure 5. E21LE matrix of evaluation research 'purpose' and 'needs' (after Mark, Henry, & Julnes, 1999)*

## SITUATING EVALUATION METHODOLOGICALLY WITHIN E21LE

The matrix above (Figure 5) allows E21LE to position the myriad existing (and upcoming) learning environment evaluation tools in a way that allows those from education, design and other stakeholders to implement evaluation strategies attuned to their purposes and needs. This theoretically-based, methodologically flexible structure has the capacity to 'house' all perceivable evaluation devices desired by multi-disciplinary teams wishing to undertake complex evaluations of learning environments.

*Snapshots* is beginning the process of populating this matrix. Each chapter describes an approach to learning environment evaluation that serves a differing set of purposes and needs. For example, the three E21LE doctoral projects have three approaches – Oliver's expert elicitation methodology (Chapter 8) serves an 'auditing' (appraisal) purpose through classification; Byer's (Chapter 9) method serves an 'improvement' purpose through identifying causality (judgement); and Sala-Oviedo's (Chapter 10) approach has a theory generation purpose (analysis) serving a strong values need. Without this type of structure, existing and upcoming evaluation tools run the risk of not meeting the complex needs and purposes evident in the sophisticated world of modern education and learning environment design.

## THE NEED FOR EVIDENCE ABOUT 'WHAT WORKS' IN LEARNING ENVIRONMENTS

It is sometimes asserted that if we had sufficient compelling evidence which argues that new generation learning environments are improving learning outcomes then there may be greater interest in transforming teaching practice, an approach occurring in the health sector where evidence-based design (EBD) has now reached a mature stage (EDAC, n.d.; HERD, n.d.). While the concept of translational research and evidence-based practice and design may be well-accepted in medicine, the models for transforming pedagogy in both school and tertiary-level teaching are still seen as much more contestable. Nevertheless, Hargreaves and Fullan (2013) argue that transforming teaching requires building professional capital, a process that is far more complex than data driven models of building business capital. Leadership for transformative change in teaching will be, they say, "a judicious mixture of push, pull, and nudge" (p. 39). The E21LE project hopes to influence the push, pull and nudge factors of pedagogical change through developing frameworks and strategies for evaluation that align practice and space.

There are two common purposes in educational evaluation, which at times are in conflict with one another. Educational institutions often require evaluation data to (1) demonstrate various forms of effectiveness to funders and other stakeholders, (2) provide a measure of performance for marketing purposes and (3) to inform evidence-based policy development. Evaluation in this context is also a professional activity that individual educators may undertake if they intend to review and

enhance the learning they are endeavouring to facilitate. Yet, the use of evaluation to drive transformative change in education is highly vexed, particularly in the higher education sector where universities value academic freedom and professional development is largely carried out through conferences and peer-to-peer networks. Any form of top-down organised transformation is hotly contested and indeed commonly resisted or corrupted. To a degree, this is true in schools as well, as teacher professional development is often left to the individual and there is often little compunction for teachers to change the way they practice.

Evaluations can be industry or academe lead. In the realm of evaluating learning environments this can promote evaluations that have a high orientation to objective/technical aspects (such as post occupancy evaluation in architecture) or those that have a high orientation to abstract/qualitative aspects (such as measures of learning outcomes in education). Certainly, previous approaches to post occupancy evaluations of learning spaces have been less concerned with pedagogy and more focussed on issues related to indoor environment quality, construction and building quality. Conversely, what is often evaluated within pedagogical practice is not only quite varied, but contested in terms of what practices are most highly valued, and rarely if ever do these evaluations cover the places and spaces for learning.

Cleveland and Fisher (2014) concluded that "approaches to evaluations that attempt to assess the effectiveness of physical learning environments in supporting pedagogical activities are in their infancy and require further development" (p. 26). They also concluded that research in this field "could profit from an interdisciplinary approach that involves people from a variety of backgrounds, including but not limited to education, human geography, environmental psychology and architecture" (p. 26).

Further to this, Lee and Tan (2011) highlighted that "while there has been much attention to the design of learning spaces over recent years, evaluations of learning spaces have been limited in depth, rigour and theoretical grounding, and heavily reliant on informal or anecdotal evidence" (p. 3). They emphasised that evaluations are highly contextual with a tendency for studies to describe outcomes positively.

The E21LE approach, developed and outlined in the pages of this book, recognises the existence of many effective evaluation strategies while also being aware of the field's evident weakness in utilising these in 'real world' circumstances. Some significant learning environment evaluation work has been, and is currently being, undertaken at the OECDs CELE (OECD, 2016a). Over the past decade there has been significant pressure for this Centre (formerly known as the Programme on Educational Buildings) to develop stronger statistical outputs from its activities, in keeping with overall OECD educational indicators work (OECD, 2015). Most recently, CELE has collaborated with CERI (the Centre for Educational Research and Innovation) in the Innovative Learning Environments project in which numbers of case study schools were evaluated in a number of countries internationally (OECD, n.d.a.; OECD, n.d.b.). CELE is now embarking on an international evaluation study of school learning environments through the PISA (Australian Council for

Educational Research, 2015) survey portal in 2015 and is proposing to survey many thousands of students and teachers as to their views on their learning environments (OECD, 2016b). In part this is being managed by ACER in Melbourne (Australian Council for Educational Research, 2016). Another international approach to evaluations is being undertaken by Educause (Educause, 2016a). This organisation has established an evaluation pilot tool – Learning Spaces Rating System – which is currently being trialled before being fully launched (Educase, 2016b). It is an excellent attempt at codifying evaluation categories so that we can make better comparisons between the impact of various learning environments. This study is USA centred but it will be a significant tool to test in other countries. The issues are similar to those encountered in Building Research Establishment Environmental Assessment Methodology (BREEAM), Leadership in Energy and Environmental Design (LEED) and Greenstar tools in terms of each of those individual country's foci. The E21LE project is, in effect, a meta-synthesis of large and small evaluations presented in a structure that allows the widest array of clients to design, implement and access cutting edge evaluations. *Snapshots* provides a glimpse of this plethora of issues and approaches.

## WHAT IS THE STRUCTURE OF *SNAPSHOTS?*

*Snapshots* is presented in three sections. Co-editor Imms introduces the *emerging issues* section, which serves the purpose of providing a glimpse of the broad range of topics that any useful evaluation matrix of learning environment evaluations must embrace. It illustrates the need, as presented in the structure of the E21LE matrix, to move beyond simplistic 'causality' evaluations into topics such as policy, teacher change, teaching practice and multi-sector evaluations, and provides a sense of the increasing complexity of this field. Next, co-editor Cleveland presents the section on *emerging methodologies*. This section illustrates the development of 'new' approaches to learning environment evaluations being developed, trialled and refined by some of our brightest minds. It illustrates how sophisticated the evaluation needs of 21st century education have become. Finally, co-editor Fisher introduces the third section, that of *emerging knowledge*. It is a tempting and challenging glimpse of the new information being generated as we speak. It serves to not only help us understand how quickly ideas are being turned into workable evaluation methods, it also highlights the significant gap that exists in the development of a significant scholarly body of knowledge as to how learning environments impact on learning outcomes.

The Chapters in this book we hope will contribute to the global works-in-progress which all aim to close this gap.

## WHO CONTRIBUTES TO *SNAPSHOTS?*

In keeping with the multi-disciplinary problem of evaluating the design and pedagogic impact of NGLEs, the chapters presented in this book come from authors

spanning a range of disciplines. Their bibliographies are presented at the conclusion of this book. These contributors are practicing architects, teachers from the tertiary sector through to primary schools, acousticians, environment space designers, and educational consultants. All have, however, two things in common. Each is a highly skilled and experienced practitioner in their own field, and each is a current, or a recently completed doctoral student undertaking research on this broad topic. Doctoral students are challenged to explore the 'dark edges' to a problem, and as such offer rigorous but 'boundary-free' thinking, facilitated by the mandate to push known knowledge to its limits. As such, no better team could undertake the work of conceptualising and trialling approaches to our new-age issue – how to evaluate the impact of the modern learning space.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge the significant contribution to sections of this chapter by Graeme Oliver, Ana Sala Oviedo, and Terry Byers.

## REFERENCES

Alkin, M., & Christie, C. (2004). An evaluation theory tree. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage. doi:http://dx.doi.org/10.4135/9781412984157

Astbury, B. (2012). Using theory in criminal justice evaluation. In E. Bowden & S. Brown (Eds.), *Perspectives on evaluating criminal justice and corrections (Advances in Program Evaluation, Volume 13)* (pp. 2–27). Bingley, UK: Emerald Group Publishing. doi:http://dx.doi.org/10.1108/S1474-7863(2012)0000013005

Australian Council for Educational Research. (2015). *Programme for International Student Assessment (PISA) Australia*. Retrieved from https://www.acer.edu.au/ozpisa/pisa-2012

Australian Council for Educational Research. (2016). *Programme for International Student Assessment (PISA) Australia*. Retrieved from https://www.acer.edu.au/ozpisa/

Campbell, D. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*, 297–312.

Carden, F., & Alkin, M. (2012). Evaluation roots: An international perspective. *Journal of Multidisciplinary Evaluation, 8*(17), 102–118.

Chelimsky, E. (2013). Balancing evaluation theory and practice in the real world. *American Journal of Evaluation*, *34*(1), 91–98.

Cleveland, B. (2011). *Engaging spaces: Innovative learning environments, pedagogies and student engagement in the middle years of school* (Unpublished PhD). The University of Melbourne, Melbourne.

Cleveland, B., & Fisher, K. (2014). The evaluation of physical learning environments: A critical review of the literature. *Learning Environments Research, 17*, 1–28. doi:10.1007/s10984-013-9149-3

Dewey, J. (1916/1997). *Democracy and education*. New York, NY: Free Press.

Dovey, K., & Fisher, K. (2014). Designing for adaptation: The school as socio-spatial assemblage. *The Journal of Architecture, 19*(1), 43–63. doi:10.1080/13602365.2014.882376

EDAC. (n.d.). *Evidence-based design accreditation and certification.* Retrieved from https://www.healthdesign.org/certification-outreach/edac

Educase. (2016a). Retrieved from http://www.educause.edu/

Educase. (2016b). *Learning space rating system*. Retrieved from http://www.educause.edu/eli/initiatives/learning-space-rating-system

Hargreaves, A., & Fullan, M. (2013). The power of professional capital. *Journal of Staff Development*, *34*(3), 36–39.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta analyses relating to achievement*. New York, NY: Routledge.

Hattie, J. (2011). *Visible learning for teachers.* New York, NY: Routledge.

HERD. (n.d.). *Health Environments Research & Design Journal*.

Imms, W. (2016). New generation learning environments: How do we know if what works is working? In W. Imms, B. Cleveland, & K. Fisher (Eds.), *Evaluating learning environments.* Rotterdam, The Netherlands: Sense Publishers.

Imms, W., Cleveland, B., Mitcheltree, H., & Fisher, B. (Eds.). (2015). *Terrains: Mapping learning environment evaluation across the design and education landscape*. Conference proceedings for Terrains: An international symposium for research higher degree research, Learning Environments Applied Research Network, Melbourne.

Lackney, J. A. (2008). Teacher environmental competence in elementary school environments. *Children, Youth and Environments, 18*(2), 133–159.

Lee, N., & Tan, S. (2011). *A comprehensive learning space evaluation model: Final report 2011*. Sydney: The Australian Learning and Teaching Council.

Love, A. J. (2010). Understanding approaches to evaluation. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., Vol. 3, pp. 798–807). Oxford: Elsevier Ltd.

Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an integrative framework for evaluation practice. *American Journal of Evaluation*, *20*(2), 177–198.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, *66*(1), 60–92. Retrieved from http://www.sfu.ca/~decaste/newlondon.htm

OECD. (2015). *Education at a glance 2015: OECD indicators*. Paris: OECD Publishing. doi:http://dx.doi.org/10.1787/eag-2015-en

OECD. (2016a). *Centre for Effective Learning Environments (CELE)*. Retrieved from http://www.oecd.org/edu/innovation-education/centreforeffectivelearningenvironmentscele/

OECD. (2016b). *Centre for effective learning environments*. Retrieved from http://www.oecd.org/edu/innovation-education/centreforeffectivelearningenvironmentscele/joinleepandsharegoodpracticeineffectivelearningenvironments.htm

OECD. (n.d.a). *Innovative learning environments*. Retrieved from http://www.oecd.org/edu/ceri/innovativelearningenvironments.htm

OECD. (n.d.b). *Innovative learning environments: The innovative cases strand*. Retrieved from http://www.oecd.org/edu/ceri/innovativecases.htm

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousands Oaks, CA: Sage Publications Inc.

Scriven, M. (1967). The methodology of evaluation. In R. Stake (Ed.), *Curriculum evaluation* (American Educational Research Association Monograph Series on Evaluation No.1). Chicago, IL: Rand McNally.

Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluation*. Thousand Oaks, CA: Sage.

Soccio, P., & Cleveland, B. (2015). *Study 1: Towards effective learning environments: An evidence based approach* (Unpublished research report). The University of Melbourne, Melbourne, VIC.

Soja, E. (1996). *Thirdspace: Expanding the geographical imagination*. Cambridge, MA: Blackwell.

Stufflebeam, D. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. S. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 117–141). Boston, MA: Kluwer-Nijhoff.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies*. Upper Saddle River, NJ: Prentice Hall.

*Wesley Imms*
*Melbourne Graduate School of Education*
*The University of Melbourne*

*Benjamin Cleveland*
*Melbourne School of Design*
*The University of Melbourne*

*Kenn Fisher*
*Melbourne School of Design*
*The University of Melbourne*