

Bridging the edtech evidence gap

A realist evaluation framework refined for complex technology initiatives

Melanie Rose Nova King

*Centre for Engineering and Design Education, Loughborough University,
Loughborough, UK*

Steve J. Rothberg

*School of Mechanical and Manufacturing Engineering,
Loughborough University, Loughborough, UK, and*

Ray J. Dawson and Firat Batmaz

*Department of Computer Science, Loughborough University,
Loughborough, UK*

18

Received 28 June 2015
Accepted 19 November 2015

Abstract

Purpose – There are five factors acting as a barrier to the effective evaluation of educational technology (edtech), which are as follows: premature timing, inappropriate techniques, rapid change, complexity of context and inconsistent terminology. The purpose of this paper is to identify new evaluation approaches that will address these and reflect on the evaluation imperative for complex technology initiatives.

Approach – An initial investigation of traditional evaluative approaches used within the technology domain was broadened to investigate the evaluation practices within social and public policy domains. Realist evaluation, a branch of theory-based evaluation, was identified and reviewed in detail. The realist approach was then refined, proposing two additional necessary steps to support mapping the technical complexity of initiatives.

Findings – A refined illustrative example of a realist evaluation framework is presented, including two novel architectural edtech domain reference models to support mapping.

Practical implications – Recommendations include building individual evaluator capacity; adopting the realist framework; the use of architectural edtech domain reference models; phased evaluation to first build theories in technology “context” and then iteratively during complex implementation chains; and community contribution to a shared map of technical and organisational complexity.

Originality – This paper makes a novel contribution by arguing the imperative for a theory-based realist approach to help redefine evaluative thinking within the IT and complex system domain. It becomes an innovative proposal with the addition of two domain reference models that tailor the approach for edtech. Its widespread adoption will help build a shared evidence base that synthesizes and surfaces “what works, for whom, in which contexts and why”, benefiting educators, IT managers, funders, policymakers and future learners.

Keywords Complex technology initiatives, Edtech domain reference models, Evaluation methodologies, Ex-ante evaluation

Paper type General Review



1. Introduction

1.1 *The evaluation imperative for educational technology in higher education*

Within higher education (HE), a number of significant factors are putting unprecedented pressure on an institution's ability to develop and invest in educational technology. The financial pressure resulting from the rapid and critical demise in funding with HE in England (a real terms cut of 46 per cent in the funding allocation between 2010/2011 and 2014/2015) (IPPR [Commission on the Future of Higher Education, 2013](#)) has meant that the institutions are faced with difficult decisions on priorities for investments and cuts. The drive towards efficiency is promoting programmes that adopt lean approaches to operational effectiveness using cost savings as an evidence of success with no guidance on evaluating any impact on the overall quality in learning and teaching (Universities UK, 2011). The migration of technology development skills away from HE into the commercial sector is another significant factor, with the number 1 challenge faced by institutions being the lack of support staff with specialist skills (UCISA, 2014).

The growing scholarly critique surfacing the distrust of educational technology (Selwyn, 2014b) and the need to take a more critical perspective on the use of technology in education (Bulfin *et al.*, 2015) compounds the uncertainty around what works, which might explain the seeming lack of appetite for a new investment in ICT. The prominence of digital systems in all aspects of HE also makes for an increasingly complex and problematic landscape of data structures and work processes across all boundaries of operations, teaching and research (Selwyn, 2014a), leading to “[deep] rooted concerns over the social, political and cultural roles of these systems”. This can be seen in dispirited accounts from in-depth interviews with academics in Australia (Hil, 2012). The use of digital technologies in general, from email to online learning systems, featured prominently as exemplifying the worst aspects of working within modern universities.

With tighter budgets, smaller teams, distrust of educational technology and disillusioned staff, it is imperative for institutions that invest in both in-house educational technology development and off-the-shelf products to not only ask “has this made a difference (in time or money saved)?” but also try and understand exactly what works, for whom and why.

1.2 *Traditional approaches to evaluation*

For the purposes of this paper, the term “edtech” is used to describe software, systems and devices that are used in HE to support the business of teaching and learning. Evaluation approaches used for edtech commonly have used formative or summative approaches that either focus on the “technology”, the “pedagogy”, the “project” or “programme initiatives”. Evaluation activities, in general, have been classified within four categories, each having their own uses (Stufflebeam and Shinkfield, 2007). *Formative evaluations* are used to provide information for developing a service, ensuring its quality or improving a particular method or approach by providing continuous feedback loops for a project. This type of evaluation is carried out before or during the implementation stage and is aimed directly at the project staff. *Summative evaluations* are retrospective and used to provide accountability reports when a product is finished or completion of a project or programme of work is attained. They are useful for determining the accountability for success or failure – aimed predominantly at sponsors or consumers. *Evaluations to assist in choice selection* are used to share proven

practices or products to help consumers make wise adoption or purchasing decisions, for example, comparisons between proprietary and open technologies with similar features (Udas and Feldstein, 2006). *Evaluations to foster enlightenment* are conducted to bring a new understanding arising from revelations. They concern themselves primarily with “why it works” by identifying the theory behind the programme. Findings from these evaluations can address particular research, theory or policy questions.

1.3 Current barriers to the effective evaluation of edtech

Formative and summative approaches are demonstrated in the review of the UK’s Joint Information System Committee (JISC) (Wilson, 2011) and the evaluations of the Centres for Excellence in Teaching and Learning (CETL). The CETL programme was HEFCE’s (Higher Education Funding Council for England) largest ever single funding initiative in teaching and learning. The inconclusive findings of both are due in part to the lack of robust evidence provided regarding the programme’s impact on teaching and learning at both the institution and sector level. This is significant because of the implications for evidence-based policy and, therefore, government funding of future edtech programmes. For example, “Only a handful of CETLs have provided evidence of the direct impact that technology-enhanced learning has had on its students, but in all cases, the belief has been that it has had a tangibly beneficial impact on learners’ and ‘Several CETLs feel that innovation in teaching and learning is being sustained, although this is not always straightforward to evidence” (HEFCE, 2011).

Five factors have been identified as a barrier to the effective evaluation of edtech (King *et al.*, 2014), these being:

- (1) *Premature timing*: Summative evaluations (of products, projects or process) carried out immediately after an edtech development will never fully give an understanding of the potential influence and impact of the initiative on learning and teaching, as it cannot take into account long-term effects.
- (2) *Inappropriate existing software evaluation techniques and models*: Existing maturity models do not help us to fully understand the complexity of organizational factors that affect the potential for success of in-house edtech development. Existing technology acceptance models are unhelpful in unearthing the complexity of staff and students’ beliefs, attitudes and intentions with regards to adopting new edtech.
- (3) *Political context and the corporatization of HE*: HE is in such a rapid state of change that it makes contextual qualitative evaluations problematic with political drivers calling for quantifiable evidence of cost savings and efficiency.
- (4) *The iterative nature of agile development and participatory design*: Homegrown edtech development is a complex process of organic and *ad hoc* product improvement.
- (5) *The semantics of edtech*: The use of inconsistent terminology within HE, often locally adapted or country specific, is a barrier to effective evaluation.

The United Nations has joined the International Evaluation Partnership Initiative (EvalPartners) and designated 2015 as the International Year of Evaluation (Rugg, 2013) to advocate and promote evaluation- and evidence-based policymaking at the

international, regional, national and local levels. A networked global multi-stakeholder process has been initiated to identify the key areas of a global evaluation agenda for 2016-2020. One of the four preliminary priorities identified so far is to strengthen individual evaluator capacity development, including the promotion of innovation of theory and new tools (EvalPartners, 2015).

In this “International Year of Evaluation”, it is timely to herald in a new innovative evaluative approach for edtech. The goal of this paper is to advocate the novel use of a particular theory-driven evaluative approach, namely realist evaluation. A realist approach will help foster enlightenment on the impact of the development and use of edtech in HE and provide “revelations” of what works. It is proposed that a shift to a theory-driven approach could help address these five factors acting as a barrier to effective evaluation (timing; technique; rapid change; complexity; and terminology). The long-term goal is to help the HE community synthesize and surface “what works, for whom, in which contexts and why” benefiting educators, funders, policymakers and, most importantly, future learners.

1.4 Objectives

The objectives of this paper are to provide a methodological review of realist evaluation and realist synthesis (already established within the health-care and social policy sectors). It proposes the innovative and novel application of realist evaluation within the domain of educational technology by articulating a realist evaluation framework specifically tailored to edtech. The authors have refined the realist approach by proposing two additional necessary steps to support the mapping of technically complex initiatives. Evaluators are also provided with two industry reference models created particularly for the classification of technology domains and associated roles that people play in relation to edtech initiatives within HE. An illustrative example is provided, describing the stage-by-stage application of the framework and reference models. A reflection on the findings and approaches taken in a recent sector review is given along with the recommendations on the practical use of the realist evaluation framework and where the future evaluative efforts should be focused.

2. Methodological review of realist evaluation

The European Commission (European Commission, 2013) has used the term theory-based impact evaluation (TBIE) to reflect a number of theory-oriented evaluation approaches developed by a number of evaluation experts (Suchman, 1967; Chen and Rossi, 1980; Weiss, 1995; Pawson and Tilley, 1997; Rogers, 2008). Theory-driven evaluation within education itself is relatively new and very rarely used for the evaluation of technology. A systematic review of the use of 45 cases of programme theory-driven evaluation approaches used between 1990 and 2009. Coryn *et al.* (2011) shows that the greatest number (47 per cent) were broadly classified as health interventions and only 1 out of the 45 was specifically concerned with a technology initiative, investigating the impact of computerized information systems on nurses’ clinical practice (Oroviogioicoechea and Watson, 2009).

2.1 What is realist evaluation?

The term “realist evaluation”, a branch of theory-based evaluation specifically for the evaluation of complex social interventions, was drawn from Pawson and Tilley’s seminal book (Pawson & Tilley, 1997). A realist approach assumes that nothing works everywhere for everyone and that context really makes a difference. It is a way of

thinking that adopts the scientific philosophy of scientific realism (Bhaskar, 1978) to uncover the underlying mechanisms and their contexts that produce distinct outcomes.

Realist evaluation begins by clarifying the “programme theory” and the mechanisms (m) that are likely to operate, the contexts (c) within which they operate and outcomes (o) that can be observed. The initial idea, the goal, the expectation, hypothesis or “programme theory” is that if certain resources (whether material, social or cognitive) are provided, then they will edge into a subject’s reasoning, generating a change in thought or behaviour. These theories (hypotheses) provide the realist evaluation with its starting point, the programme theory being the unit of analysis rather than the programme itself. Theories are generated and evidence is, then, collected in the form of context (c) + mechanism (m) = outcome (o) configurations in sentence-like configurations C + M = O called CMOCs (pronounced *seemocs*) in the realist literature. These are then analysed and form a starting point of “the intervention theory works under conditions X, Y and Z” as an if-then proposition (Pawson and Sridharan, 2010).

2.1.1 The evaluator as theorizer. The action of realist theorizing is based on a method of thinking called retroduction (also known as abductive reasoning), a logic of inquiry also found within scientific realism. This means relying on your previous expertise, experiences, hunches or imagination to generate a theory that is *inspired* by the evidence. As a realist evaluator, you are in fact a theorizer using retroduction with a combination of deduction (theory *tested* against evidence) and induction (theory *derived* through evidence) with an element of inspired and creative thinking.

Generating theories may require a workshop involving other evaluators, commissioners, programme and policy staff. However, it may also require looking outside the actual intervention itself and examining similar interventions in other policy areas to identify for whom, where and how they appeared to work to, therefore, help generate a theory. Generating and refining theories can also involve extrapolation from formal theories in a similar theory domain, for example, theories on technology adoption or incentivisation. Pawson terms the role of these formal theories within realist evaluation as “re-usable conceptual platforms” that help evaluators build on lessons learnt from previous programmes that shared a similar component theory (Pawson, 2013). The rationale being that an evaluation should never start from scratch and must build on lessons from evaluations in the past.

2.1.2 Programme mechanisms. Finding programme mechanisms is fundamental to theorizing how and why programmes work within realist evaluation. Programme mechanisms are participants’ reaction (change in beliefs, desires and behaviour) to the mixture of the resources made available to them by the programme. Mechanisms have three main characteristics:

- (1) mechanisms are usually hidden;
- (2) mechanisms are sensitive to variations in context; and
- (3) mechanisms generate outcomes (Astbury and Leeuw, 2010).

2.1.3 A realist’s approach to complexity. A basic assumption of realist evaluation is that “programmes are complex interventions introduced into complex social systems”. Pawson (2013) provides realist evaluators with a checklist (Table I) for identifying the key characteristics of the complexity of the programme under the acronym VICTORE (Volitions, Implementation, Contexts, Time, Outcomes, Rivalry, Emergence). The point

Table I.
The VICTORE
complexity checklist

<i>Volitions</i>	The particular choices made by the programme subjects
<i>Implementation</i>	The linked and disparate components in the chain of implementing a policy and programme
<i>Context</i>	The individuals, inter-personal relations, institutional settings and the wider social, economic and cultural setting of the programme
<i>Time</i>	The timing of policies and the sequencing of programmes
<i>Outcomes</i>	The planned (and unplanned) activities under the programme and their outcomes
<i>Rivalry</i>	Interventions are always dispensed into a world of existing interventions
<i>Emergence</i>	Components in a system will often combine to produce novel components, thus changing the composition of the system under investigation

of this is for evaluators to step back before commencing the design of the evaluation research to first map the complexity landscape of a programme, which therefore helps them to focus on the relevant areas of complex systems and purposefully take a limited cut at specific issues.

2.2 What is realist synthesis?

Realist synthesis (or realist review) is a realist’s approach to systematic review. It is a secondary approach that applies a realist philosophy to the synthesis of findings from primary studies that have a bearing on a single research question or set of questions. For example, reviewers may begin by drawing out from the literature the main ideas that went into the making of certain types of interventions (the programme theory). This programme theory hypothesizes how and why a class of intervention is thought to “work” to generate the outcome(s) of interest. The theories are then tested using relevant evidence (qualitative, quantitative, comparative, administrative and so on) from the primary literature on that type of intervention. Realist synthesis can also be used to help answer a current policy question about a proposed initiative. Realist synthesis of existing literature can also be used within a current evaluation to help generate candidate theories and test them. The first application of realist synthesis within the edtech domain was published in 2010 for online medical education (Wong *et al.*, 2010).

2.3 When is realist evaluation appropriate?

Realist evaluation is appropriate if (Westhorp, 2014) the initiative is new or a trial or pilot programme that seems to work but “for whom and why” is not yet understood. It is appropriate for evaluating interventions that will be scaled up for more users to understand how to adapt the intervention for new contexts and for evaluating programmes that have previously demonstrated mixed patterns of outcomes to understand why the differences occur. The method is also suitable for the ex-ante evaluation of a policy or programme idea.

Realist evaluation is not appropriate if there is no particular new initiative or process that has been introduced under investigation or if there is not enough time or resource available to undertake a realist approach (see Section 4.3 for recommendations). If the only requirement is to find out if the initiative made a difference to a clearly defined objective but one does not need to know why, then other evaluative methods will be more appropriate.

3. A realist evaluation framework refined for complex edtech initiatives: an illustrated example

3.1 The realist evaluation framework

The following example is not intended to provide detailed guidance on how to conduct a realist evaluation; for this, we direct interested readers to summary articles or various publications on the methods (Pawson and Tilley, 1997; Westhorp, 2014; Astbury and Leeuw, 2010; Dalkin *et al.*, 2015). However, it provides a useful visual and conceptual framework for those new to realist evaluation, outlining the stages and necessary steps within the evaluation life cycle (Figure 1), which are as follows: preparation, mapping, theory formation, abstraction, the cycle of evaluative enquiry and presentation of findings. The framework has been created to make explicit for edtech evaluators the process of realist evaluation and, therefore, expedite its transition and adoption from predominantly health-care settings to education. More importantly, the mapping stage

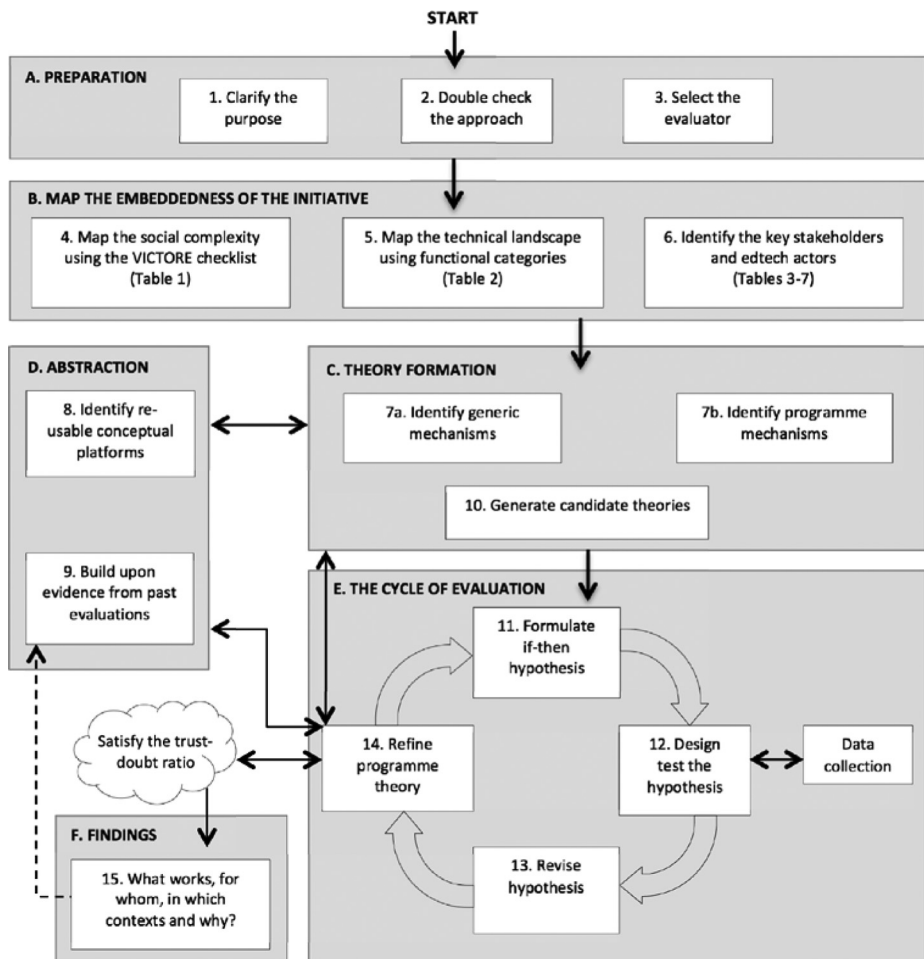


Figure 1. The realist evaluation framework refined for complex edtech initiatives

has been refined with two reference models (the addition of Steps 5 and 6) to address the mapping of complex technology initiatives within education. The framework is underpinned by the organizing principles of evaluation science, as set out in Pawson's realist manifesto "The Science of Evaluation" (Pawson, 2013), which has provided a blueprint for realist evaluation as a scientific discipline.

3.1.1 Refining our framework to address technology complexity. One of the factors impinging on effective edtech evaluation is the use of inconsistent *terminology* used for describing technology and edtech roles in education. With regard to the "Context" element within the VICTORE complexity checklist, we have refined this mapping stage by the inclusion of Steps 5 and 6, which are particular to edtech. This provides evaluators with two architectural reference models as a way to classify and categorize the technologies in their functional domain as well as the actors involved in the initiative to aid in the investigation of context across a myriad of possible organizational set-ups, technical architectures and varying job titles.

3.1.2 An illustrative example. To aid in the transition of the realist approach to edtech evaluation, it is proposed that the general term "initiative" should be used rather than programme. An initiative is defined as "an act or strategy intended to resolve a difficulty or improve a situation; a fresh approach to something". The range of edtech initiatives suitable for an investigation could conceptually cover a multitude of potential *things*, such as projects, institutional strategy or policies, a new process adopted (e.g. co-design), a new online course, new interactive set of teaching material, in-house software development, deployment of new commercial software or devices.

Sections 3.2-3.7 provide a descriptive overview of each stage within the framework and illustrate certain steps using the following fictional scenario: A few departments within an institution have implemented the same automated attendance monitoring initiative but with mixed outcomes.

3.2 Stage A: preparation

The realist evaluation is intended to inform institutional policy and practice; therefore, collaboration with strategic leaders is needed to fully understand the purpose of the evaluation and to reveal "how will the answers be used?" For example, to answer an institutional policy question:

Should automated attendance monitoring be mandated for all departments for all taught sessions when we do not know why it seems to work in Department X but not Department Y?

After checking the approach (realist evaluation is appropriate because it is already known that the initiative is successful in Department X, but it is not known why exactly and the ambition is to scale up the initiative), the evaluator sets out to gain the widest possible understanding of the array of possible influences that shape the fortunes of the initiative in each department. They draw on their knowledge of the educational technology literature (synthesis), their own technical expertise and that of others, and most importantly, they do not ignore hunches based on their past experiences.

Upfront work, on the generation of hypotheses to test, may be required if it is problematic to uncover the initial program theories, in other words, the rationale of why the programme was expected to work by the creators. Additional work will also be needed if there are no previous evaluations that give clues about what might be affecting

whether and how the initiative works or the evaluators are not up to speed with the relevant edtech literature. Preparation work could include investing in a preliminary research project to develop realist program theory that can be used as the basis for multiple evaluations in the future.

If there are no resources to invest in necessary preliminary work, then it is possible to construct realist evaluations to be theory building rather than theory testing. It implies a heavier focus on qualitative work to investigate mechanisms and how they are affected by the context. A staged design, over a few evaluations, will mean doing more qualitative work in the first one or two evaluations to develop the theories and, then, more quantitative and mixed method evaluations later to test these theories across types of technologies, types of curricula or student cohorts or other important features of context that emerge.

3.3 Stage B: mapping the embeddedness of the initiative

3.3.1 Step 4: mapping the complexity. Initially, a rough mapping of the initiative, using the VICTORE complexity checklist, is carried out. During the evaluative cycle and data collection, the complexity landscape gains more detail about programme mechanisms and variations in context (Table II).

3.3.2 Step 5: mapping the technical landscape using the edtech functional domain reference model. It is imperative that evaluators are specific and consistent when defining technology type during the “context mapping” stage of the VICTORE

Volitions	What were the reactions of staff and students to the choices made available to them, as a part of the attendance monitoring initiative?
Implementation	Has the attendance monitoring initiative been implemented differently in Department X? What does take-up look like with different types of staff members? How long did it take to implement? Did the software and systems work in the anticipated way? Were they actually usable?
Context	What support and resources had been given to each department? What local and institutional policies were in place? What choices were available to staff and students regarding use? What are the previous experiences of staff and students in relation to attendance monitoring? How successfully has the technology been implemented? Was it reliable and trusted to work every time?
Time	When was the initiative brought in and announced? Did it affect other workload commitments? Did the use of it take longer than anticipated?
Outcomes	What planned (and unplanned) activities happened during the roll-out of the initiative? What were the tangible outcomes both positive and negative? What was the quantifiable difference in attendance at sessions? Did it make a difference to staff and student behaviour and attitudes?
Rivalry	What other initiatives were being rolled out at the same time? For example, personal tutoring systems or peer-mentoring programmes. What alternative edtech was available that fulfilled the functional requirements for attendance monitoring, for example, spreadsheets and iPads?
Emergence	What impact have the other initiatives had to produce novel components of the attendance monitoring initiative? Thus, has the original initiative under investigation now emerged as a holistic engagement monitoring initiative within Department X, for example?

Table II.
Potential areas for investigation to aid the mapping of the attendance monitoring initiative

complexity checklist. We propose the addition of this distinct next step, particular to the realist evaluation of edtech, which refines the approach and provides the essential technical nuance to enable a common classification and understanding of technological type and, therefore, the purpose within the initiative. The technological context of an institution can be metaphorically thought of as “the digital campus”, which can be an umbrella term for the Information and Communication Technology (ICT) infrastructure that a member of teaching staff, staff involved in supporting teaching and learning and the students themselves will need to interact with as a part of their time at the University – the ICT infrastructure being composed of the hardware and computing devices, web applications, software and the data itself. A deliberately broad set of technology domains has been included in the edtech functional domain reference model (Table III), as conceivably any technology that a student interacts with will have an impact on their overall student experience, but each technological domain performs a particular and distinct function within the digital campus.

For this scenario, a mapping of the primary and secondary functions of the technologies that are used within the attendance monitoring initiative can be classified as shown in Table IV.

3.3.3 Step 6: identifying key stakeholders using the edtech actor domain reference model. It is also imperative that evaluators are specific and consistent when defining particular edtech-related roles during the “context mapping” stage of the VICTORE complexity checklist. We therefore propose the addition of this distinct next step in the framework. Tables V-IX outlines an edtech actor domain model (abstract roles or actual jobs). As a part of the mapping, it is important to identify certain roles that might be expected to make an initiative work based on formal theories, for example, a “technology evangelist” (Table IX), even if that role is not fulfilled in practice, as its absence could have an impact on the outcomes.

For our scenario of an attendance monitoring initiative, a mapping of the people involved in the initiative within the institution might be as shown in Table X.

3.4 Stage C: theory formation

The initial programme theory is derived, for example, from asking the Head of Department X why they thought the automated attendance-monitoring initiative would work (Table XI).

Using retroduction, the evaluator begins thinking about the component resources of the initiative (based on the initial mapping) and people’s potential reactions (programme mechanisms) to them that might be triggered within different contexts. Theories are crafted using $C + M = O$ configurations.

3.5 Stage D: abstraction

Underlying generic mechanisms and candidate theories can also be generated by drawing on formal theories in the literature (Step 8) and from synthesizing findings from previous evaluations (Step 9), thus abstracting away from looking solely at the initiative itself.

3.5.1 Step 8: Identifying potentially useful re-usable conceptual platforms for edtech. Understanding the generic mechanisms (behaviours and thoughts) that are triggered by the resources provided by an initiative means fundamentally understanding human

Functional domains	Definition
Classroom technologies	Hardware, web applications and software used by teachers within the digital classroom to help deliver teaching and assessments as well as manage courses as a whole <i>For example, Course delivery platforms such as the Virtual Learning Environment (VLE) or Massive Open Online Course (MOOC) platform technologies. Lecture capture, plagiarism detection and curriculum software</i>
Content creation tools	Devices, software and web applications used to create digital content for any purpose, including collaborative content. For example, text-based content, images, animations, videos, audio etc. <i>For example, Digital cameras, MS Office; Adobe Creative Suite; Google Docs; Camtasia</i>
Personal space and persona technologies	Hardware, web applications and software used by both learners and teachers. Helping learners to take control and manage their own learning as well as providing the tools for development and promotion outside the curriculum. Usually seen as personally owned technologies <i>For example, E-portfolios, drop-in computer labs, informal collaboration tables, mobile devices</i>
Research tools	Hardware, web applications and software that support any kind of research activity undertaken by staff and students <i>For example, SPSS; NVivo, Bristol Online Survey, EndNote, RefWorks, high-performance computing (HPC), electronic lab books</i>
Relationship management systems	Systems that manage and enhance the relationships between staff and students, the organization and industry, the organization and potential learners. These technologies help manage contacts and communications by incorporating relationship workflow management and are client (learner, company) centered <i>For example, Student Enrolment customer relationship management systems (CRM), fundraising and donor management systems</i>
Data warehouse	A system whose primary function is to collect and store raw data that underpins one or more applications <i>For example, Student information systems</i>
Repositories and knowledge management systems	A system whose primary function is to catalogue, manage access to and retrieval from and provide search and discovery functionality to digital assets, for example, remote folders, digital library content, papers <i>For example, Open access repository for publications (DSpace), shared workspaces, reading lists management system (RLMS), Library management system (Aleph), Library catalogue plus (LCP)</i>
Dashboards for business intelligence and analytics	A system that collects and/or utilizes data from people and other applications and is used to provide analytical insights into the organization as a whole <i>For example, Management applications (data extracts, management reporting), online module feedback systems. Engagement monitoring tools</i>
Communication technologies	Hardware, web applications, software and APIs that support digital communication of any kind <i>For example, Email, chat clients, web and video conferencing, telephony, Skype, license servers</i>
Operational, infrastructure and access technologies	Hardware, web applications, software and APIs that support the physical asset management, corporate operations and institutional obligations as well as managing access to something or someone <i>For example, Authentication, network and middleware, Antivirus, IT help desk, Timetabling system, room-booking systems, research asset management, CAD software asset management, HR, corporate website CMS, Finance, accommodation system, campus and building access, technology-equipped lecture rooms</i>

Table III.
A proposed edtech functional domain reference model

behaviour and why people react in certain ways. For example, one might look at the work of Michie *et al.*, who have created the behaviour change wheel (BCW) as a new method for characterizing and designing behaviour change interventions. It provides a useful conceptual platform for identifying contexts and associated behavioural mechanisms (Michie *et al.*, 2011). Social network analysis is the use of network theory to analyze social networks. In the context of education and technology, emerging research of network formation within online communities provides a conceptual platform for understanding the mechanisms at play within these types of virtual contexts (Groenewegen and Moser, 2014). Models that help explain people’s adoption or acceptance of technology (Venkatesh *et al.*, 2003) can be used as a basis for theory formation with regards to mechanisms.

Teaching and learning theories provide a wealth of opportunity to help candidate theory formation by providing conceptual platforms to understand learner and teacher behaviour. For example, in the context of open education and Massive Open Online Courses (MOOCs), the theory of rhizomatic learning provides a model for the construction of knowledge in an unbounded and exploratory way by participants.

	A web application that creates registers for sessions and stores attendance data	Data capture devices; networked and portable	Learner management system	Table IV. Technologies identified during mapping and classified within primary and secondary definitions
Primary	Classroom technology	Infrastructure and access technology	Business intelligence	
Secondary	Data warehouse		Communications	

Functional domain lead	The person(s) who is an expert in the functioning of software and tools in that domain of technology, requiring specialist expertise and knowledge. For example, a domain lead for “classroom technologies” would require a technical expert who also had expertise in pedagogy. This role may contribute to policy and procedure related to their domain across schools and committees. Might take a lead project role in new technology initiatives within their domain			Table V. Roles linked to a particular specialist technology
Functional domain specialist	The person(s) who actively develops, maintains or supports technology in a particular domain. Often is on hand to resolve more specialist technical issues and horizon scans within their technical domain			

Problem poser	The person(s) who has identified and articulated the problem that needs to be solved. Provides a real-world candidate use case for a potential solution			Table VI. Roles linked to technology pre-evaluation, selection and procurement
Policy protagonist	The originator of the plan or strategy (solution) for undertaking the initiative or the lead advocate or champion of the particular reform (policy) proposed			
Technical gatekeeper	“Serves as an intermediary between the average member of the firm and external sources of information” (Allen, 1977). Has a significant impact on the innovation process and technology development of an organization. Highly experienced, intuitive and exerts a high level of influence within the organization (Scheiner <i>et al.</i> , 2014)			

Table VII.
Roles linked to in-house technology development

Problem analyst	The person who is working with the problem poser and documenting and communicating the problem to the development team
Solution architect	The person who designs and shares the vision of the solution, whether a combination of technology, systems, data, policies, documentation, training, roll out and embedding activities
Product developer Interaction designer	The software engineer, programmer or web developer of the product Designs how people will interact with the digital product. Could also design the interface elements of an application. Within edtech, this could be the person who is designing the learning pathways through applications
Content producer	Produces content for an application. For edtech applications, this is often the subject expert who may be working with a learning technologist to transpose existing paper-based teaching material into a digital (perhaps interactive) format
DevOps	A blanket term for systems engineers, system administrators, operations staff, release engineers, database administrators, network engineers, security professionals and various other sub-disciplines. Patrick Debois first coined the term DevOps in 2009. (The Agile Admin, 2010)
Systems architect	Ensures that the digital product connects and integrates seamlessly within the system and data infrastructure of the digital campus

Table VIII.
Roles linked to technology implementation

Capability owner	The person who is strategically responsible for the institution to have the capabilities that this technology brings, this might be delegated downwards. The capability owner might oversee a portfolio of technologies, across all functional domains, which provide capabilities within their own area of responsibility. Ultimately, this might be Pro-Vice Chancellors, Chief Operating Officers, Registrar or Director of Finance
Product owner	The custodian is the internal champion and guardian of the system or software on behalf of the capability owner. It is the responsibility of the custodian to ensure that the technology has appropriate technical management and a service is in place for users to implement the technology to its full capability. The custodian is also responsible for reviewing functionality and whether the technology still provides the capability required and prioritizing development based on business value. The Product Owner holds the ongoing vision for the product
Technical manager	If the system breaks, then the person who is technically responsible for sorting it out

Fittingly, a useful guide is provided by the Open University in their open education platform OpenLearn (The Open University).

Conceptual platforms also help to provide a scaffold within which to map the complexity of context, particularly with regard to the organisational context. For example, organisational development theory explains how organisational structures and processes influence worker behaviour and motivation. There are also many maturity models for mapping the complexity of organisational contexts, for example, within e-learning ([Marshall, 2010](#)) and domain-specific models, such as the student engagement

Service manager	The person responsible for implementing the technology into staff and students' practices, making sure end users have training and support in place to use the technology effectively, encouraging maximum usage and diagnosing user issues. This person is the "face" of the technology
Educational support provider	Provides assistance to the primary user, usually teachers, to enable the effective use of the product within the education setting. This role can be carried out by designated support staff or by colleagues, for example, within community groups
Technical support provider	People within the organization that usually provide Tier 1, 2 and 3 technical support. The supplier or vendor provides Tier 4 (Walker, 2001)
Primary end user	The user (e.g. the teacher) who has selected or is in charge of "driving" or "providing" the technology to the secondary user. Their use is more active in shaping the technology, and there is a consequence if the primary user does not use the technology to its full capability. Primary end users can also display specific attributes of use, for example, the Diffusion of Innovation (Rogers, 2003) types of innovators, early adopters, early majority and late majority
Secondary user	The passive user of a technology, usually a participant of the technology and on the receiving end of the product once a primary user, has intervened. This is usually the student or learner
Indirectly affected	Those people who are indirectly affected by the use of the technology by others but who are not primary or secondary users. For example, if a teacher asks everyone in their class to tweet their thoughts on the current topic of discussion, then displays back the syndicated feed to the class, the students who did not participate for any reason are those indirectly affected
Technology Evangelist	"a person who builds a critical mass of support for a given technology, and then establishes it as a technical standard in a market that is subject to network effects" (Lucas-Conwell, 2006)

Table IX.
Roles linked to technology use and adoption

Department X	Department Y	University
Educational support provider (Department E-Learning officer)	Primary end users (Lecturers) Secondary users (Students) Secondary users (Personal Tutors)	Functional domain lead for access technologies (Head of Networks) Policy protagonist (Head of Department X)
Technical support provider (Local IT)	Technology evangelist (First-Year Tutor)	Technical gatekeeper (Enterprise Architect)
Primary end users (Lecturers)		Capability owner (PVC-Teaching and Director of IT Services)
Secondary users (Students)		Technical manager (Head of Student Systems)
Indirectly affected (Department Office Staff)		Technical support provider (IT Services)

Table X.
Participants identified during mapping of contexts and classified by edtech role definition (actual job title)

success and retention maturity model (SESR-MM), which “will indicate the capability of HEIs to manage and improve SESR programs and strategies” (Clarke *et al.*, 2013).

It is usual to select only one or two formal theories that are directly relevant to the priority question of the investigation. For this scenario, the evaluator can create candidate theories that focus on the “context” of the functional domain of the technology and capability it brings. For example, “Access Technologies” and “Business intelligence” bring the capability of knowing where each student is at key points, once they have signalled electronically their presence in the classroom (Table XII).

Or the evaluator could choose to focus on “mechanisms” at play, for example, the psychological concept of trust (or lack of) and empathy between teachers and students (Table XIII).

Table XI.
Initial programme theory

Context	+ Mechanism	= Outcome
Students in their first year in HE have previously been used to their attendance being taken in class in FE and at school. They think this is the norm for “lessons”	The act of signing in with their student card, from Day 1, reinforces the idea that it is expected and important that they turn up to every session	Students display a consistent pattern of attendance at those classes where registers are taken

Table XII.
Context-based conceptual platform: post-panopticism in high-technology human tracking systems and the notion of power relationships

Context	+ Mechanism	= Outcome
Department X chose to use generic system emails that get sent to non-attenders automatically	Students receiving a non-personalized email from a “system” added to the sense of being anonymously surveilled by “the institution”	Students were more likely to communicate reasons for non-attendance with non-face-to-face methods, such as emails or department forms

Source: Dobson and Fisher (2007)

Table XIII.
Mechanism-based conceptual platform: theory of experiential learning and trust relationships between teacher and student

Context	+ Mechanism	= Outcome
The personal tutor within Department Y received an email alert about a significant drop in attendance for one of their tutees. They arranged a face-to-face meeting with the student	The personal tutor demonstrated empathy when they spoke to the student. In turn, the student felt they could trust their tutor to support them when they needed more help and encouragement	The student felt more inclined to attend lectures, even though they found the topics challenging

Source: Kolb (2015)

3.6 Stage E: the cycle of evaluation

The evaluator then performs a cycle of enquiry based on the candidate theories generated. If-then hypotheses are formulated (Step 11), and then, evidence is collected that supports, rejects or refines them (Step 12), ensuring that the data collection focuses on evidence that can refine the candidate theories. Based on the initial round of investigation, hypotheses are revised (Step 13), and the evaluator's theories are refined when significant CMO patterns emerge (Step 14). If there is a long list of potential theories to test, then the evaluator could choose to use a panel of experts to help sift and sort them into a priority list, for example, by using the Delphi technique (Hsu and Sandford, 2007) for gaining a consensus of opinion on the focus of the investigation.

The emerging CMO models can be grouped into themes to help focus the next iterative cycle of investigation. For example, theories $C1 + M1 = O1$ and $C1 + M2 = O2$ work when the department provides a personal tutoring system to the students as an intervention to encourage attendance. The models can also be sequential, for example, within an implementation chain, if theory $C1 + M1 = O1$ is present, then the outcome from the first implementation chain (O1) goes on to provide the context for the next part of the intervention; therefore, $C2(O1) + M2 = O2$.

The nature of a complex initiative means there could be an ever-increasing number of theories to test and refine. The task of the evaluator is to focus their enquiries on refining theories that answer the questions specifically posed by the evaluation commissioners and trust that they are focusing on the right slice of the pie. The evaluator needs to trust certain assumptions they make about the programme and the nature of the evidence they have collected that confirms their hypothesis. If they are in doubt, then Steps 11 to 14 are repeated in a cycle of evaluative inquiry. Ray Pawson terms this as the trust-doubt ratio (Pawson, 2013).

3.7 Stage F: findings

Findings are presented in linked CMO configurations that tell the story of the intervention. Some evaluators prefer to do this in prose form, others in tabulated formats. The findings need to communicate the essence of "why did it work?" and whether the focus was on refining the theories within particular contexts or refining within particular mechanisms (thoughts and behaviours) that were triggered as a result of the initiative and the outcomes that were found as a result. Initial candidate theories that the evaluator did not have time to test or those that were untestable can still be included in the findings. This is helpful for the decision-makers to consider these theories when deciding on new initiatives in the future. The evaluation takes only a small cut at the complexity of the initiative under investigation, but future evaluation in this area can build upon the theories generated in an attempt to cumulate knowledge about why initiatives work or not, for whom and in which contexts.

4. Discussion

4.1 The significance of this new approach to kick start a necessary change in evaluative thinking

This paper argues the imperative for a theory-based realist approach to help redefine evaluative thinking within the IT and complex system domain. Realist evaluation and

realist synthesis are an advanced approach to evaluation research, but they offer an innovative and insightful means by which to further our understanding of complex social and technological interventions in HE and their resulting impact on the staff and student experience. Presenting findings from evaluations in a realist way will help policymakers and funders to connect emotionally to the evidence of why an initiative has worked or not, as it unearths and communicates the complex human stories as well as shining a spotlight on a small area of the complex social and technical systems within which we all work and live.

The rationale of advocating a realist evaluation approach is also in its potential to address four of the five factors impinging on effective evaluation of edtech (*timing; technique; rapid change; and complexity*). The *timing* issue, addressed as a theory-based evaluation, can be conducted before, during or after implementation, as it is testing the underlying assumptions of an edtech initiative. The emphasis on evaluating the social system within and around an intervention will mitigate against the narrow conventional *techniques* of evaluating solely the software or specific technology-based models to understand usage and adoption. The VICTORE mapping takes into account the *complexity* inherent in people, organisations, rival interventions and *rapid change*, as well as the complex development (often agile) implementation chains of in-house edtech development. Most importantly, it will provide evidence to support a “why” hypothesis for specific sub-groups and contexts. With current government and institutional drivers for evidence of “did it make a difference?” (usually using financial or performance indicators) (Universities UK, 2011), there is now a growing need for funders to know exactly “why it made it difference” in certain contexts and within complex programmes that can only be answered using a realist approach.

The final factor *terminology* is addressed by the use of the edtech functional domain model (a component of “context” or a “resource” element within an initiative) and the edtech actors domain model (the “who”), newly proposed in this paper. This will help to theme the context element of the CMO candidate theories and surface any semi-predictable patterns emerging. This is significant not only for edtech evaluators but also for the HE sector as a whole. The adoption of these reference models by evaluators will provide a way of consistently communicating findings relating to the role played by particular technologies and people within edtech initiatives. Therefore, this makes it easier to synthesize findings from disparate investigations and builds an evidence base for the sector, helping to cumulate knowledge for evidence-based policymaking and funding decisions in the future.

4.2 Cumulative realist learning

Realist evaluation should both learn from and build upon previous investigations and contribute findings to shared evidence bases. Therefore, it is worthwhile to briefly reflect on a recent report (Trowler *et al.*, 2014) the findings from which not only justify the need for a sector-wide adoption of the realist approach to edtech evaluations (their research used theory-driven but not explicitly realist evaluation) but also give us clues of where to prioritize future evaluative efforts.

The evaluation team at Lancaster adopted a theory of change perspective to review previous evaluative evidence and 15 key stakeholders’ opinions of HEFCE-funded teaching and learning enhancement initiatives from 2005 to 2012.

Although technology-enhanced learning was not specifically in scope, the review considered HEFCE-initiated enhancement activities, which encompass edtech, including the Centres for Excellence in Teaching and Learning (CETLs), as well as evidence from sector surveys, such as the UK's National Student Survey (NSS).

The theoretical framework, underpinning the investigation, provides future realist evaluators with some good re-usable conceptual platforms for further investigations. Enhancement initiatives were classified within a selection of theories of change (e.g. contagion from good examples; technological determinism; rewards and sanctions; consumer empowerment; professional imperative) as well as the educational ideological positions of stakeholders (traditionalism; progressivism; enterprise and social reconstructionism). The framework also describes some strategic aims of particular interventions, which will help realist evaluators in theorizing outcomes (e.g. increased efficiency, increase equity of experience, change in teaching and learning practice, change in power relations between students and staff).

A key finding from the stakeholder interviews was the “need for better data about enhancement requirements, prioritization of efforts and good evaluation of outcomes and effects”. In fact, the perceived lack of good evaluation is a consistent theme throughout most of the initiatives discussed. Perhaps it was, therefore, presumptive for these stakeholders to also assert that “large, high-profile projects often do not represent good value for money” (Trowler *et al.*, 2014, p. 3) without perceived “good” evaluation in place. It may be that the key finding, in relation to this sentiment, should be that stakeholder preferences for meaningful evaluative outcomes, from high-profile projects, is for quantifiable evidence of cost (i.e. financial) benefits over qualitative evidence of long-term effects.

The team acknowledges the limited resources allocated to the investigation by the Higher Education Academy (HEA), so it is not surprising that findings are generalized, lacking detail about positive outcomes in particular contexts and tended to report stakeholder perceptions of impact at the institutional or sector level. This in itself is significant as the report acknowledges:

[...] much depended on the situated circumstances of the intervention at institutional level.
[...] the combined effects of the strategy tended to have different outcomes in different locales and made it difficult to determine whole system effects (Trowler *et al.*, 2014, p. 8).

The answer to how a whole system “changes” in response to local innovations is something that eludes the sector. From the interviews and wider study, weakness was acknowledged as: scaling up to the sector-level change; lack of understanding of the change processes or theories of change; and change only associated with early adopters or niche practitioners. The team warns, however, on the implication of these findings for effective policymaking within such complex-adaptive systems:

It is always tempting to make decisions based on a technical-rational understanding of change-processes. However, we know that micro-political and macro-political processes as well as the robust defence of turf, careers, reputations and position mean that change is more often a process of “muddling through” in a loosely coupled way than a rational process of successive goal setting and achievement (Trowler *et al.*, 2014, p. 26).

There are two points for realist evaluators to take away here. There is work to do in eliciting the rationale of the policymakers in “why they expected the initiative to work”. Deciphering their views as a theory of change, that is, stages in the process of change (or the implementation chain), that underpins their vision of events that will unfold. Each stage in the process will subtly affect the context and outcomes of the next phase of implementation and therefore, participants’ reaction (mechanisms) to them. Successive adaptations, in fact, change the conditions that may have made the initiative work in the first place. For example, there appears to be a tipping point reached when “local” or “innovative” activities cease to become effective. Understanding what it is about the context and the mechanisms at play, at that tipping point, will perhaps help policymakers unlock the key to widespread and systematic changes in practice and culture and be able to evidence this in a rigorous way. The second point for evaluators to take away is that this method can also be used to evaluate the robustness and likelihood of success of future policy initiatives, which has to be more cost-effective than initiating investments without any foresight.

Any details about geographical and organisational context, as well as technical capability (or digital literacy) of the key stakeholders, was missing from the investigation itself. An intentional but significant omission was exploring the link between outcomes and the rapidly changing technological context over the period 2005-2012. The inability to technically scale up initiatives may well have had a deleterious effect on the perception of the initiative as a whole. However, the team recognizes that “a constellation of factors is shifting the HE section into unfamiliar territory. [...] (that) represents a powerful new vector”. These (e.g. technological change, globalization) provide significant “contexts”, within which to focus future evaluative efforts.

4.3 Recommendations

Based on the methodological review of using realist evaluation and our illustrative example, we suggest the following recommendations for the focused and practical use of the edtech realist evaluation framework:

- (1) *Build evaluator capacity* to quickly get up to speed with the realist approach to evaluation, the newly established Centre for Advancement in Realist Evaluation and Synthesis (CARES) at the University of Liverpool (UK), runs regularly events and an international conference for budding realist evaluators. Help and support is available from the open-access RAMESES (Realist And Meta-narrative Evidence Synthesis: Evolving Standards) online discussion list (RAMESES, 2015).
- (2) *Be clear on the purpose* of the evaluation and policy question being answered to ensure that the realist approach is appropriate. Invest time in designing and planning the evaluation thoroughly with the resources you have available.
- (3) *Use the edtech architectural domain models* to classify both the technology function and the actors as a part of the evaluation. This will be key in transferring lessons from one evaluation to others and, therefore, synthesize “what works, for whom, in which contexts and why?”

-
- (5) *Adopt the RAMESES realist quality standards* when conducting and reporting on realist syntheses and/or meta-narrative reviews (Wong *et al.*, 2013).
 - (6) *Evaluative efforts* should be directed into:
 - Conducting realist synthesis of existing primary evidence, from the edtech literature, for the purpose of realist theory building particularly in “context”-driven theory creation, for example, technological context.
 - Unearthing the contexts, mechanisms and outcomes within stages of complex implementation chains to shed light on the causality of change at the macro level.
 - Contributing to a shared map of complexity with regard to the shifting HE landscape.

5. Conclusions

This paper provides a comprehensive review of the realist approach and innovatively proposes its application within the domain of edtech. We provide an edtech realist evaluation framework, refined for complex technology initiatives by providing evaluators with a novel taxonomy of technology types and actors within the edtech domain. The addition of these refinements will make it much easier to synthesize findings from disparate edtech investigations on what works, for whom, in which contexts and why enabling cumulative realist learning in the sector. An illustrative example is provided, specifically for edtech, not available elsewhere in the literature. The example shows how the framework can be applied to undertake a complex evaluation of an automated attendance monitoring initiative. It demonstrates its potential to uncover the human, technical and organisational factors that affect this often contentious policy within HE. An argument is presented for how this refined realist approach will have the potential to address the five factors currently impinging on the effective evaluation of edtech (*timing; technique; rapid change; complexity and terminology*). A reflection on the findings and approaches taken in a recent sector review is provided, along with the recommendations for the focused and practical use of realist evaluation, for those interested in conducting evaluations of technology in use within teaching and learning or for the evaluation of future policy initiatives.

References

- Allen, T.J. (1977), *Managing the Flow of Technology*, MIT Press, Cambridge.
- Astbury, B. and Leeuw, F.L. (2010), “Unpacking black boxes: mechanisms and theory building in evaluation”, *American Journal of Evaluation*, Vol. 31 No. 3, pp. 363-381.
- Bhaskar, R. (1978), *A Realist Theory of Science*, Routledge, London.
- Bulfin, S., Johnson, F.N. and Bigum, C. (Eds) (2015), *Critical Perspectives on Technology and Education*, Palgrave Macmillan, Basingstoke.
- Chen, H.T. and Rossi, P.H. (1980), “The multi-goal, theory-driven approach to evaluation: a model linking basic and applied social science”, *Social Forces*, Vol. 59 No. 1, pp. 106-122.
- Clarke, J.A., Nelson, K.J. and Stoodley, I.D. (2013), “The place of higher education institutions in assessing student engagement, success and retention: a maturity model to guide practice”,

Higher Education Research and Development Society of Australasia, AUT University, Auckland.

- Coryn, C.L., Noakes, L.A., Westine, C.D. and Schröter, D.C. (2011), "A systematic review of theory-driven evaluation practice from 1990 to 2009", *American Journal of Evaluation*, Vol. 32 No. 2, pp. 199-226.
- Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B. and Lhussier, M. (2015), "What's in a mechanism? Development of a key concept in realist evaluation", *Implementation Science*, Vol. 10 No. 49.
- Dobson, J.E. and Fisher, P.F. (2007), "The panopticon's changing geography", *The Geographical Review*, Vol. 97 No. 3, pp. 307-323.
- European Commission. (2013), *Evalsed Sourcebook: Methods and Techniques*, European Commission, Regional Policy, Brussels.
- EvalPartners (2015), "Global Evaluation Agenda 2016-2020 Preliminary outcomes of the global networked online consultation, from EvalPartners", available at: http://mymande.org/sites/default/files/files/Global_Evaluation_Agenda2016-2020_v3.pdf (accessed 6 January 2015).
- Groenewegen, P. and Moser, C. (2014), "Online communities: challenges and opportunities for social network research", in Brass, D.J., Labianca, G., Mehra, A., Halgin, D.S. and Borgatti, S.P. (Eds), *Research in the Sociology of Organizations (Contemporary Perspectives on Organizational Social Networks)*, Emerald Publishing, Bingley, Vol. 40, pp. 464-477.
- HEFCE (2011), "Summative evaluation of the CETL programme – Final report by SQW to HEFCE and DEL", HEFCE.
- Hil, R. (2012), *Whackademia: An Insider's Account of the Troubled University*, UNSW Press, New South Wales.
- Hsu, C.C. and Sandford, B.A. (2007), "The Delphi technique: making sense of consensus", *Practical Assessment, Research & Evaluation*, Vol. 12 No. 10.
- IPPR Commission on the Future of Higher Education. (2013), *A Critical Path: Securing the Future of Higher Education*, The Institute for Public Policy Research, IPPR, London.
- King, M., Dawson, R., Batmaz, F. and Rothberg, S. (2014), "The need for evidence innovation in educational technology evaluation", in Uhomoibi, J. (Ed.), *Proceedings of INSPIRE XIX: Global Issues in IT Education*, British Computer Society, Southampton, pp. 9-23.
- Kolb, D.A. (2015), *Experiential Learning: Experience as the Source of Learning and Development*, in Neidlinger, A. (Ed.), 2nd ed., Pearson Education Inc, New Jersey.
- Lucas-Conwell, F. (2006), "Technology evangelists: a leadership survey, from Growth Resources", available at: www.gri.co/pub/res/pdf/TechEvangelist.pdf (accessed 3 August 2015).
- Marshall, S. (2010), "A quality framework for continuous improvement of e-learning: the e-learning maturity model", *Journal of Distance Education*, Vol. 24 No. 1, pp. 143-166.
- Michie, S., Stralen, M.M. and West, R. (2011), "The behaviour change wheel: a new method for characterising and designing behaviour change interventions", *Implementation Science*, Vol. 6 No. 42.
- Nuffield Department of Primary Care Health Sciences (2015), "Mailing list", The RAMESES Projects, available at: www.ramesesproject.org/index.php?pr=Mailing_list (accessed 8 February 2016).

-
- Oroviogioicochea, C. and Watson, R. (2009), "A quantitative analysis of the impact of a computerized information system on nurses' clinical practice using a realistic evaluation framework", *International Journal of Medical Informatics*, Vol. 78 No. 12, pp. 839-849.
- Pawson, R. (2013), *The Science of Evaluation: A Realist Manifesto*, Sage, London.
- Pawson, R. and Sridharan, S. (2010), "Theory-driven evaluation of public health programmes", in Killoran, A. and Kelly, A. (Eds), *Evidence Based Public Health*, Oxford University Press, Oxford.
- Pawson, R. and Tilley, N. (1997), *Realist Evaluation*, Sage, London.
- Rogers, E.M. (2003), *Diffusion of Innovations*, 5th ed., Free Press, New York, NY.
- Rogers, P.J. (2008), "Using programme theory to evaluate complicated and complex aspects of interventions", *Evaluation*, Vol. 14 No. 1, pp. 29-48.
- Rugg, D. (2013), May. "UNEG joins EvalPartners in declaring 2015 as international year of evaluation, from United Nations Evaluation Group (UNEG)", available at: www.unevaluation.org/document/detail/1378 (accessed 6 January 2015).
- Scheiner, C.W., Baccarella, C.V., Bessant, J. and Voigt, K.I. (2014), "Thinking patterns and gut feeling in technology identification and evaluation", *Technological Forecasting & Social Change*, Vol. 12 No. 1.
- Selwyn, N. (2014a), *Digital Technology and the Contemporary University*, Routledge, London.
- Selwyn, N. (2014b), *Distrusting Educational Technology: Critical Questions for Changing Times*, Routledge, London.
- Stufflebeam, D.L. and Shinkfield, A.J. (2007), *Evaluation Theory, Models, & Application*, Jossey-Bass, San Francisco.
- Suchman, E. (1967), *Evaluative Research*, Russell Sage Foundation, New York, NY.
- The Agile Admin (2010), "What is DevOps?", The Agile Admin, available at: <http://theagileadmin.com/what-is-devops/> (accessed 25 February 2015).
- Trowler, P., Ashwin, P. and Saunders, M. (2014), *The Role of HEFCE in Teaching and Learning Enhancement: A Review of Evaluative Evidence*, Lancaster University, Centre for Higher Education Research and Evaluation, Higher Education Academy, Lancaster.
- UCISA (2014), *2014 Survey of Technology Enhanced Learning for Higher Education in the UK*, Universities and Colleges Information Systems Association, UCISA, Oxford.
- Udas, K. and Feldstein, M. (2006), *Apples to Apples: Guidelines for Comparative Evaluation of Proprietary and Open Educational Technology Systems*, State University of New York, SUNY Learning Network, The Observatory on Borderless Higher Education, London.
- Universities UK (2011), *Efficiency and Effectiveness in Higher Education*, Universities UK, London.
- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology: toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.
- Walker, G. (2001), *IT Problem Management*, Prentice Hall, New Jersey.
- Weiss, C.H. (1995), "Nothing as practical as good theory: exploring theory-based evaluation for comprehensive community initiatives for children and families", in Connell, J.I., Kubisch, A.C., Schorr, L.B. and Weiss, C.H. (Eds), *New Approaches to Evaluating Community Initiatives*, The Aspen Institute, Washington DC, pp. 65-92.
- Westhorp, G. (2014), *Realist Impact Evaluation: An Introduction*, Overseas Development Institute (ODI), Methods Lab, ODI, London.

- Wilson, A. (2011), *Review of the Joint Information Systems Committee*, HEFCE, London.
- Wong, G., Greenhalgh, T. and Pawson, R. (2010), "Internet-based medical education: a realist review of what works, for whom and in what circumstances", *BMC Medical Education*, Vol. 10 No. 12.
- Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J. and Pawson, R. (2013), "RAMESES Publication standards: realist syntheses", *BMC Medicine*, Vol. 11 No. 21.

Further reading

The Open University (n.d.), "Open Education – 5.5 Rhizomatic Learning", OpenLearn, available at: www.open.edu/openlearn/education/open-education/content-section-7.5 (accessed 8 March 2015).

Corresponding author

Melanie Rose Nova King can be contacted at: m.r.n.king@lboro.ac.uk