Book chapter submitted to:

Learning Environments: Technologies, Challenges and Impact Assessment

Title:

An integrated educational technology evaluation framework: en route to a generic, iterative model for establishing learning effectiveness.

submitted by:

Stephen Quinton
[BEd (Curtin); MEd (UWA); PhD (UWA)]
Email: straqu@gmail.com

An integrated educational technology evaluation framework: en route to a generic, iterative model for establishing learning effectiveness.

ABSTRACT

The introduction of computer-dependent technologies from the early 1990s into university and school curricula precipitated an unprecedented revolution in the way learning was designed and delivered. Today, a plethora of technology offerings affords the opportunity to cater for the many diverse preferences and needs that students expect to be integral to their day-to-day learning experiences.

The factors vital to ensuring successful technology adoption within higher education and discussed at length by researchers over the past two decades have changed markedly with the rapid expansion in the range of technology choices that are now available to students and lecturers anywhere, anytime. In a market saturated with educational technology and software options, the current abundance of technology choices are at once exciting, challenging, complex, and by implication, replete with problems for the unwary. This development has naturally magnified the complexity of comparing the educational value and effectiveness of online learning with face-to-face learning.

The University of New South Wales (UNSW) has established a major organisational-wide programme to evaluate and integrate educational technologies into its learning and teaching offerings. This programme is referred to as the Technology Enabled Learning and Teaching (TELT) Platform. One result of this effort has been the establishment of a generic educational technology evaluation framework that has successfully underpinned the university's technology adoption process from first contact with software vendors through to the determination of whether a given technology is educationally effective and therefore approved for use by staff and students throughout the campus.

A one year review of the research literature on institutional technology adoption and related evaluation studies highlighted a distinct gap in understanding the interdependences between the organisational structure and its processes, and the needs and issues of faculties and schools. Such understanding is essential for establishing a durable educational technology implementation strategy that meets the full spectrum of institutional, faculty, and school expectations. This review led to the conclusion that there is a strong imperative for universities to establish an ongoing organisational-wide evaluation programme that assists to make valid technology choices and underpins the compilation of evidence that informs the effective integration of technologies into their learning and teaching provisions. The evidence that has been derived to date at UNSW has ensured continued improvement in the quality of the university's online learning and teaching programmes.

Central to the construction of an effective educational technology evaluation strategy is the task of strengthening inter-departmental cohesion from which to cultivate a common awareness of how technology can support learning and teaching, whilst preserving the ideologically divergent yet essential pedagogical, technological, and business imperatives of the university's facilities, schools and discipline areas. However, for a technology evaluation strategy to be truly effective, such common understanding must also incorporate the many unique interests, issues, and needs of faculties and schools that when considered as a whole, ensure broad acceptance of online learning. This is where the TELT Evaluation Framework comes into play.

This chapter presents an overview of the research that led to the establishment of the TELT Evaluation Framework followed by a description of its main components and an explanation of how its integral iterative cycles align with the university's strategies and plans to improve the quality of its technology supported learning and teaching programmes.

INTRODUCTION

In April 2009, the Director of Learning and Teaching at the University of New South Wales (UNSW) commissioned a review of the university's Technologies Enhancing Learning and Teaching (TELT) environment. This environment comprises the information and communication technologies that support teaching and learning in each of the on-campus, distance, and blended modes and more broadly, it applies to the institutional setting in which these technologies are used. It encompasses the university's learning management systems (LMS), other software applications used either in association with or separately from the LMS, the institutional planning, policy, and governance processes that guide and direct the learning and teaching and management processes that in turn, support the provision of services and quality assurance as well as the training and support processes for students and staff.

The aim of the review was to provide a snapshot of the current status of the TELT environment, and more specifically to (Technologies Enabled Learning and Teaching (TELT), 2009, p. 2):

- facilitate key players (that is, those with specific responsibilities for providing direction and leadership over the TELT environment) in undertaking a self-assessment of the state of the TELT environment
- analyse the elements of the TELT environment and their inter-relationships, highlighting strengths and weaknesses
- help inform the University's future planning of its TELT environment, at both strategic
- and operational levels, and
- provide baseline data against which future comparisons can be made to monitor changes and improvements.

The report presented a comprehensive baseline that demarked the issues surrounding the successful rollout of the TELT Platform and mapped out the critical areas to be targeted for improvement. Of equal importance, the report indicated which organisational structures are missing and need to be in place to ensure the successful implementation of educational technologies across the university.

The overarching finding of the report is that decision-making in the past has occurred without reference to valid evidence-based evaluation processes. This omission was viewed as a serious concern as indicated by the realisation that many sections of the report contain links to evaluation related issues. The other major concerns raised in the report included: the perceived lack of clarity around the roles and responsibilities with regard to TELT; and, the absence of clarity due to unclear policy, quality assurance, and management procedures. As will be explained later in this chapter, Stage Three of the TELT Evaluation Framework (see the section titled "The Stages and Sub-layers of the TELT Evaluation Model") now addresses TELT quality assurance through a logically constructed evaluation cycle, that now informs and provides clarity on policy and management related decisions.

In recognition of the issues and problems identified in the report, the Learning and Teaching unit at UNSW (L&T@UNSW) initiated a "change in roles and attitudes approach" in direct response to feedback received on staff reactions to past educational technology implementation programmes (Technologies Enabled Learning and Teaching (TELT), 2009). The unit shifted its traditional approach and practices to staff support where service calls were mixed with course design requests and new educational technology adoptions were dependent on top-down decisions, to a reappraisal of the roles of educational designers and the establishment of a new evaluative model for informing the approval and incorporation of technologies into the institutional infrastructure (see Figure 3). The changes in roles, attitudes, and approaches were three-fold:

1. Under the realigned structure, a significant area of change occurred in the provision of technical versus course design assistance. The Learning and Teaching unit, which previously was perceived as the first point of contact for simple 'how-to-use' technical questions, reviewed its role and

subsequently focused on the provision of educational support for technology use and integration into course design and delivery. This change was an important factor in gaining the faculties' approval and support on a series of follow-on technology-related initiatives. The uncertainty of the role of the Learning and Teaching unit in previous initiatives had generated widespread disbelieve in the educational value and usefulness of technologies in the context of a large research institution (Technologies Enabled Learning and Teaching (TELT), 2009).

2. The second area of change involved a concerted effort to mitigate staff and faculties' attitudes toward technology. As pointed out by Anderson and Johnson (2006), large research universities tend to prefer a loosely connected organisational structure with faculties operating almost as independent states within a State, each assuming high levels of autonomy. This arrangement often leads to faculty resistance to centralised decision-making processes.

   The new structure enabled full inclusion of the eclectic mix of faculty plans, views, and needs, not merely as recipients of organisational decisions about technology adoption, but as partners in the evaluation procedures that inform such decisions. This change necessitated the arduous task of overcoming the mistrust of technology use in education, a longstanding factor that often impedes the achievement of successful technology adoption programmes. Another advantage to this approach stems from the affordances it creates for establishing the necessary institutional linkages that are crucial to effecting genuine improvement in learning and teaching methods and learning strategies across the campus.

3. The third area of change applied to the inclusion of learners' voices in the evaluative process. It is often assumed that evaluation methodologies should reflect the importance of students' feedback and opinion rather than confine them to the role of end recipient of a delivered educational service, which as a consequence leads to decisions that are made without their input. However, the reality is that these assumptions do not always correspond to actual practice as in the case reported by Britain and Liber (2004). Equally vital, is the need to avoid a repetition of the example reported by Alexander and Golja (2007), where learners were formally included in the evaluation process, but their voices were used to justify the adoption of a technology solution, which overshadowed an otherwise valid attempt to improve the quality of the learning outcomes.

THE CALL FOR A NEW EVALUATIVE FRAMEWORK

The evaluation of technologies in education has a long history. Starting from industry-oriented evaluative models such as the K-12 educational models of Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, and Provus (1971), Kirkpatrick's (1983) four steps to measuring training effectiveness, and the Xerox approach to managing training systems (Alden, 1983), technology evaluation was considered primarily as a subset of educational evaluation just as the virtual campus was mistakenly viewed as a sub-branch of the main (physical) campus in the early stages of distance education expansion and experimentation.

However, similar to the manner in which the transportation revolution occurred through the invention of the automobile, educational technology turned out to be not just another tool or delivery method, but brought into play a specific set of learning and teaching principles and curriculum requirements. This unprecedented development established technology-enhanced learning and teaching as an educational domain in its own right, with its own rules and capacities to deliver a wide variety of new experiences and outcomes for learners, lecturers, and the educational institution. Although the evaluation of educational technology now spans more than two decades, for many models and implementation schemas several important elements remain problematic.

One significant outcome, namely the influence of different instructional media on the learning process, was brought to the attention of pioneer educational technology evaluators in the early 1990s. The learning ↔ technology dichotomy generated considerable debate within the educational

community with the main idea that both aspects – the structuring of the learning activities (instructional design), and the methods for delivering those activities (technology) require careful evaluation in order to determine the totality of the learning process (Clark, 1994). Simply adapting an evaluation instrument used for traditional on-campus courses does not capture all the multiple facets of evaluating a technology-enhanced course (Roberts, Irani, Lundy & Telg, 2004). The notion of a learning ↔ technology dichotomy exerted significant influence on the development of the educational technologies evaluation field and is discussed further in the section titled 'Biases'.

Taken as a whole, the optimistic predictions made by educational psychologists on how technologies are best applied to support and enhance teaching and learning (Kozma, 1994), are still being validated today. As a consequence, a wide range of questions about the effective use of technologies in education still remain unanswered. Evaluative concerns expressed a decade ago by Bruce (1999) have only been partially integrated into current models. In the section to follow, the discussion focuses on the questions of why the evaluation of educational technologies presents such a challenge for educational communities and what can be done to resolve the complex question of how to devise an all-inclusive framework for the effective evaluation of educational technologies.

There are a variety of evaluation models that are used in the educational technology field. They range from toolkits (Oliver, McBean, Conole & Harvey, 2002) to cookbooks (Learning Technology Dissemination Initiative, 1998), along with new evaluative models that are still under development such as the institutionalisation model of Nelson, Post & Bickel (2001) or the Software Process Improvement and Capability dEtermination (SPICE) model (Marshall & Mitchell, 2004). The questions of why new models are still appearing after twenty or so years of educational technology research and what are the contributing reasons raises a number of issues in reviewing the literature on educational technologies evaluation:

1. Technologies are in constant development

   Traditional formative and summative evaluation methods do not fit neatly into the eternal kaleidoscope of changing technological tools, systems, and architectures. Although formative evaluation aims to assess products and processes as they are developed, the technological development cycle never ceases. Thus, in the sense of establishing a defined end point for a completed project or product, the inherent goal of summative evaluation may not be feasible. The development process is subject to transition from one stage of its development to the next. Therefore, in the area of educational technology evaluation, there is a cogent need for an evaluation framework that accommodates the iterative evaluation processes that are characteristic of many technology-related processes, whilst as argued by McNaught (2002), simultaneously embracing the eclectic myriad of teaching and learning functions required for successful delivery through the Web.

   Within the domain of computer science, the word "iteration" has a specific meaning: any project progresses through several iterations; new tools and solutions are refined through several development-test-improvement cycles; and, the development of new tools commences with prototyping. Such iterative processes mandates the need for an evaluative framework that is also iterative in its approach to solve the "once something is completed it is already obsolete" (Bruce, 1999) dilemma encountered by many researchers with an educational, curriculum, and instructional background.

   To some extent, several evaluation models have addressed these problems. The Marshall and Mitchell SPICE model (2004) includes iterative evaluation cycles; and the Agostinho, Oliver, Harper, Hedberg, and Wills (2002) Evaluation and Redevelopment Framework (ERF) model includes possibilities for addressing technological redevelopment issues. However, both these models, one being technological, the other embracing concerns about the needs of learners, do not

address the relevance and effectiveness of the applied pedagogies, or detail the adequacy of the technological evaluation criteria.

2. Narrow scope of the existing models

   From the times when eLearning, technology-enhanced learning, and distance education were viewed as simply the other delivery option within the broader university context, the notion of educational technologies evaluation morphed into a programme of evaluation that is conducted along administrative lines and is narrowly focused on the institution's preparedness to support and promote the programme. By starting with an industry-oriented, business evaluative model, the educational technologies evaluation field could not escape one of the fallacies of a business-related approach when the scope of evaluative activities is limited to a specific delivery programme (Agostinho et al., 2002; Franklin, Yoakam, & Warren, 1996), or even with the application of a single tool (Crozat, Hû, & Trigano, 1999). According to Lefoe, Gunn, and Hedberg (2002), this tendency translates into evaluation reports that are relevant only (or even confidential) to the particular institution (or country) that initiated the study.

   In general terms, as Baker and Herman (2003) note, two prevalent approaches to technology evaluation in the past fifteen years apply to: the evaluation of specific technologies (a narrow scope), and broad, often longitudinal studies where the results become irrelevant long before they are realised due to the rapid rate of change in technology. Indeed, even models with a clearly articulated pedagogical agenda such as the Britain and Liber (2004) conversational model focus on existing Course Management Systems in their evaluative purposes, but avoid examination of broader enterprise level solutions or virtual learning environments such as Second Life (2010).

   The essential problem with a narrow evaluative scope as applied to technology evaluation is that information technologies are also communication technologies, and as a result produce a network effect (Bruce, 1999). A tool used in one classroom may be used very differently in another context; all collectively generating a relatively broad application area that may or may not be visible or relevant within the framework of one institution or a specific course programme.

3. Weak theory base and *ad-hoc* evaluations

   While early educational technology evaluation models were grounded in psychological / measurement theories, the rapid rate of change in technology generated a situation where a planned evaluation study was not completely thought through at the implementation or even the initial programme design phase. Out of necessity, people with varying, non-evaluative backgrounds conducted the evaluation study in an ad hoc manner in an attempt to assess and improve the use of a technology.

   As a result, there are a number of evaluative models that do not contain a strong rationale, a valid theoretical foundation, or a statement of purpose, and attempt to compensate for the absence of these elements by: presenting evaluators with context-specific selections of questions to submit to interviewees (Mehlenbacher, Bennett, Bird, Ivey, Lucas, Morton, & Whitman, 2005); or advancing a "just-in-time" evaluative delivery approach (Reeves & Hedberg, 2003); or preparing tailor-made evaluative plans that draw on an existing pool of questions (Lam & McNaught, 2004). Other models (such as Pangaro, 2000) produce a self-fulfilling effect in the sense that the evaluative criteria or schema are not connected to existing literature, but instead rely on internal design guidelines.

4. Prescriptive vs. descriptive evaluative models

   In general terms, evaluation models can be divided into prescriptive (most common), which specify how evaluations are conducted (exemplars); or descriptive, a set of statements and generalisations that describe, predict, or explain the evaluation activities. Prescriptive models may

include iterative evaluation cycles (Marshall and Mitchell, 2004), but in essence the iterations are imposed on evaluators and stakeholders instead of being an integral part of the evaluative activities.

Descriptive models, on the other hand, allow for an authentic reflection of the evaluation activities in real time (Cousins and Whitmore, 1998) and can be considered a viable solution for educational technology studies in light of the rapid rate of technological change as raised beforehand. Furthermore, descriptive models are participatory: the stakeholders themselves are involved in an empowering process of democratic decision-making and the active promotion of social change. These models better reflect the reactions and impressions of the participants and stakeholders. In such instances, where evaluation is conducted in a not-for-profit organisation (for example, a public educational institution), the descriptive evaluation model presents a viable alternative to the prescriptive model.

One other reason to consider a descriptive, participatory evaluation model is the fact that all the stakeholders (including students) are given control over the creation of the knowledge (Zammuto, 1982) and the informational input that is required to effect the evolution of a final practical summation.

Within the area of educational technologies evaluation, many models appear to be prescriptive, with rare exceptions (Lam & McNaught, 2004). However, even these exceptions represent a collection of evaluation-related activities rather than full-scale models, and simply offer basic background information on how to evaluate different pedagogical, technological, and organisational components.

5. Value-free versus Multiple values evaluation

   The notion of value-free evaluation is closely related to the questions surrounding evaluator-driven (prescriptive) versus stakeholders and evaluator driven, real time descriptive evaluation. The prescriptive approach inevitably generates the question of "whose values" are reflected in the particular evaluation process. One way to solve this problem is to embed multiple values in the evaluative process (Zammuto, 1982).

   Embedding multiple values in an evaluative framework allows for avoidance of the evaluator's biases as well as assists to reveal a detailed picture of the environment / technology under examination. This strategy also requires more resources and time to be allocated to the evaluative activities (Chen, 1990). However, modern pluralistic societies consider addressing the needs of different members of the evaluation audience is a necessary attribute (Stufflebeam & Shinkfield, 2007). While there is little doubt about the necessity, it is also a difficult task to resolve.

Although there is no shortage of frameworks or models for evaluating different types of technology, there is however, a marked absence of generic, inclusive, multiple-values evaluations of educational technologies that are grounded in a strong teaching and learning philosophy. The majority of models reflect the values of the evaluator, the institution, or even a software manufacturer (as highlighted by comparing the Institute for Higher Education Policy (2000) benchmarks and the eLearning Maturity Model (eMM) (Marshall, 2007). Recent evaluative models are still struggling with the multiple values challenge (Lam & McNaught, 2004) as well as the prescriptive versus descriptive challenge (Bach, 1994).

BIASES

Following directly on from the previous section, the research literature reviewed clearly confirms that the learning ↔ technology dichotomy was widely discussed by the educational technology community during the 1990s. This dichotomy also influenced the field of technology evaluation since

the evaluators themselves were either from a software engineering background, or a pedagogical background, and many were not professional evaluators. These two major backgrounds (technological and pedagogical) generated two distinct lines of inquiry in examining the purposes and the development of evaluation models. While some models espouse a technological stance (Marshall & Mitchell, 2004; Mehlenbacher, et al., 2005), other models support a pedagogical approach with a minimal attention paid to organisational and technological details (Boud & Prosser, 2002; Dean, Biner, & Coenen, 1996).

In particular, Marshall and Mitchell's SPICE model includes a learning / pedagogical component in its later version, but the component itself (that is, the questionnaire) has been updated from a study sponsored by a major educational software company (expanded on in a later section). Thus, a certain bias favouring this particular company may be present in the latest version of SPICE. Conversely, Boud and Prosser's model highlights technology features that are important for the development of learners' critical thinking abilities and do not cover the crucial organisational and technical implementation aspects of technology evaluation.

While there are also models that blend technological and pedagogical approaches, they have not escaped the challenges inherent in evaluating technological factors such as the innate structure of the application that determine interface design and navigational capacity, and the impact of these aspects on learning (Agostinho, et al., 2002). In some instances, the organisational and technological questionnaires used in these models are far more detailed than the pedagogical component (Attwell, 2006).

A substantive reason a pure technological approach to educational technology evaluation became popular in the build up to the new millennium, is the effect of technological progress on the growing "informatisation" of society, that is, the "shift of the central sources of economic activity from the primary and secondary sectors towards the services and information sectors of the economy" (Cunningham, Ryan, Stedman, Tapsall, Bagdon, Flew, & Coaldrake, 1998, p.13). Technology is in vogue, and is no longer viewed merely as a tool - it possesses integral values that may directly or indirectly influence pedagogical affordances and uses. Often, the fact that technology has its own pedagogical orientation (Britain & Liber, 1999) is completely overlooked.

Faculties are advised to move "beyond intuitive value" and "adapt to the changing landscape" (Carmean & Brown, 2005, p. 5) of new technological implementations while the reasons that make the landscape change – technological progress or the commercial interests of software developers, are still unclear. It is for this reason that viewing a LMS (for example) as a starting point for learning in online environments was widely criticised in the research literature (Siemens, 2004).

Meanwhile, one notable approach to educational technology evaluation gave rise to beneficial outcomes, namely quality assurance standards and benchmarks for technology developers and users. To some extent, quality assurance benchmarks in distance learning and education promoted the development of technology evaluation by creating structural elements on which to direct and guide the evaluation process. However, benchmarks should be developed systematically, not ad hoc. The early examples of "cherry-picking" best practices in education and devising benchmarks based on those practices failed to live up the promise (Hagner, 2001).

It is also important to distinguish evaluation benchmarks from evaluation models. By definition, benchmarks provide a fixed baseline against which to gauge the progress and effectiveness of the evaluation process as opposed to the much broader evaluation model (or framework) that assists to structure the evaluation process to explore the multiple facets of the phenomenon under examination.

What happens if the baseline is skewed (or even biased)? The Institute for Higher Education Policy benchmarks (2000) were widely adopted for technology evaluation. The benchmarks generated by this study (which included a set of Institutional support, Course development, Teaching and Learning, Course structure, Student support, Faculty support, and Evaluation and assessment guidelines) were

further adopted by Marshall for his eMM model (2007) and referenced by many educational technology researchers. The problem however, was that this set of benchmarks was originally generated by one of the largest monopolies in the educational software field (the study was commissioned by Blackboard Inc.). Although Blackboard was not a monopoly in 2000, the ability to describe and influence the major functions of their software tools during the development phases may have aided in the transition to monopoly status. The potential to guide the outcomes of software development may also prevent other important needs from being met by educational software companies, particularly if a widely accepted of benchmarks was already in place.

Then there are instances where the practices and the processes of teaching and learning are largely ignored in the pursuit of functions, features, integration, and a myriad of technological and business concerns. Thus, while quality assurance benchmarks are useful in reflecting the essential baseline, the origin of those benchmarks should also be carefully examined in order to prevent a situation where the development of the evaluation criteria is conducted by the entity that in effect, is the subject of the evaluation.

One other bias related to the issues noted above is what could be referred to as the ‘additive’ approach to technological integration. When different technology systems are integrated (such as BlackBoard and student record systems), the question of how teaching and learning will benefit is often overlooked (Boys, 2002). This situation is not simply the result of implementation issues, but is a direct consequence of whenever faculty voices are not heard or not taken into account by educational administrators. Moreover, although there are instances of benchmarks related to students, there is no evidence of attempts to refine such benchmarks to include them as vital stakeholders in the evaluation and modification of educational technology tools.

The integration of different technologies should promote the use of different learning and teaching philosophies and methods. Thus, a generic evaluative framework that facilitates evaluation of the final product or the synthesis of several technological systems may prove to be a necessary alternative to the piecemeal evaluation of different components of a complex technological mosaic. Britain and Liber (2004) also argue for a comprehensive (learner-centred) approach to evaluation as opposed to an additive (administration-centred) approach.

In their study, Britain and Liber confirmed that virtual learning environments changed markedly during the five years from 1999 to 2004 in that they focused less on teachers and administrators, and more on learners themselves. Nevertheless, they believe that inherent biases still exist since the largest educational software manufacturers still use third party products (for example Oracle) to improve the quality of their software offerings. Thus, virtual learning environments design could ultimately be considered (third party) technology driven.

In essence, there is much evidence to suggest that educational technology evaluation requires new integrative approaches that avoid the faulty Evaluator = Evaluation formula that is independent of the evaluator’s background, and covers multiple aspects of the evaluation process starting from the organisational to the pedagogical, and extending to socio-cultural factors.

It is feasible therefore, that unwanted biases could be eliminated with the formulation of a generic evaluative framework that encompasses research findings from systematically conducted educational technology evaluations. Such a generic framework can incorporate a wide range of teaching and learning practices that are not confined to one particular institution or by the outcomes of one specific sample. This framework needs to be broad in its scope to inform and improve the learning and teaching practices that are applied within and across different departments, schools, faculties, and institutions.

Last but not least, in addition to enhancing learning and teaching, the evaluation framework should also assess how well the theoretical framework itself improves technology evaluation activities and informs how best to incorporate new technology tools into teaching and learning practice. The long-

term goal of the TELT evaluation strategy is to devise an evaluation framework that encompasses the iterative processes that underpin technology development, tracks the trajectory of success across faculties, and over time enhances divergent learning and teaching approaches.

OPEN SOURCE EVALUATION

The evaluation of Open Source solutions stands out as a separate area from general (commercial) technology evaluation models. Partially, it can be explained by the origin of the Open Source software philosophy as discussed below; and partially by the desire of the commercial software companies to eliminate free educational technology options by integrating Open Source products into their own products range.

In 2005, Barbara Ross (a co-founder of WebCT) envisioned the integration of the open source and commercial tools in an attempt to support diverse pedagogies and learning outcomes (Ross, 2005). According to Ross, the main features of such integrated systems should include support for teaching, learning and research. However, this approach appeared to view open source software as no more than another piece of cake on the institutional enterprise plate. Although on the surface, the idea of integration is worthwhile and thought-provoking, no organisational or technical details followed from the vision statement made by the author of the chapter.

An alternative integrative approach was proposed by the educational software creator - ANGEL Learning. In their view, course management systems should be open to allow low-level customisation with the client institution's information technology services and needs because it is not always cost-effective for the educational software creators to embed this level of customisation (Mills, 2005). Is open source just a quick and cheap solution for the customisation problem or is it a stand-alone instructional solution? One of the best features of Open Source platforms that made them popular with Asian distance education institutions is exactly the ability to handle customisation issues in line with the needs of developing countries (Batpurev, 2005).

The preferred model for educational technologies evaluation should acknowledge the uniqueness of the conditions that brought Open Source into existence. First, according to an early theorist of the Open Source movement Eric Raymond, Open Source originally came into play out of necessity to "scratch a developer's personal itch" and the strong personal motivations of the developers involved into Open Source projects that almost guarantee high quality software solutions (Raymond, 1998). Second, software that is open to everyone, such as Facebook or a blog, tend to prompt users to cooperate and apply their efforts to the attainment of a self-regulated learning ethos. Moreover, open source software presents a less threatening learning environment for students when compared to the reduced flexibilities and restrictions of commercial products.

In a sense, commercial products emulate the traditional educational system with their inherent structures and divisions that may serve as conscientisation (Freire, 2007) tools that manoeuvre students to inadvertently absorb the boundaries and functional order that in turn may intimidate or disrupt the learning effectiveness experienced by other more independent, constructivist orientated students. By default, most commercial learning management systems afford an objective / instructivist course design approach, whereas open source solutions permit constructivist learning approaches.

In a Freirian sense, Open Source presents a more "authentic" approach to learning thus allowing students and developers to be aware of their individual learning needs as well as the shortcomings of the software and as a result, seek ways to work cooperatively to solve these issues. The collaborative opportunities afforded by Open Source software assist to invert imposed limitations and allow truly democratic concepts of education and technological literacy to arise (Stokes & Stokes, 1996).

The potential for Open Source platforms to add pedagogical / critical theory value to the learning process clearly demarks the core difference between Open Source and proprietary Course Management Systems (CMS). The latter according to Arnold (2003), appeared to have cut training

costs, and their raison d'être has a business rather than educational nature. Thus, two highly contradictory tendencies reflected in Open Source software solutions can be identified: it is software developer-driven; and at the same time, presents a less restrictive learning environment for students. It also illustrates the innate influences of applying a pedagogical versus a technologically led focus to developing (and evaluating) educational technology.

Unfortunately, most literature on Open source evaluation is limited to technological and business factors without taking the pedagogical characteristics of the software into consideration. For example, the Karin van den Berg synthesised model (2006) includes such factors as community, release, activity, longevity, license, support, documentation, security, functionality, and integration. However, her notion of community refers to software developers and technologically savvy audiences rather than educators. In a similar way, several additional factors in van den Berg's model refer to technological and business variables. A further example of a technological approach to Open source Evaluation is presented in the work of Graf and List (2005) where all pedagogical factors are covered under the single checklist category of 'user-friendliness'.

Nonetheless, despite the variety and sophistication of evaluation approaches that can be applied to Open Source software, the positive effect of Open Source concept on students' learning appears to have been unnoticed by the educational technology community.

The preceding factors and examples of the evaluative models that comprise these factors point to a disarrayed mosaic of contemporary educational technology evaluation practice: while there are models that include some of the factors mentioned above, there are also no models that include all these factors. The components that structured the development of a new evaluation framework at UNSW and the rationale for determining the framework's core design principles are briefly discussed in next section.

COMPONENTS OF A GENERIC TECHNOLOGY EVALUATION MODEL FOR THE TELT PLATFORM

Methodologies

One of the first steps toward the creation of a generic evaluation framework was made by Hedberg, Wills, Oliver, Harper and Agostinho (2002) through their ICT (Information and Communication Technologies) project. They revised the possibilities for developing generic evaluation rubrics based on high quality learning activity designs. Their model included the key elements of the learner's engagement, learning context, challenge, and practice. However, their approach, just as in the case of Britain and Liber (2004), seem to focus on a "micro-world" learning activity level and applied grounded theory methodologies where existing practices "build" on the new evaluation framework.

This approach holds some promise for existing technologies, but it neglects the iterative nature of new tools as they become available in an ever changing world of technological advancement. As with any grounded theory case, it perfectly identifies what is already in place, but does not cover crucial additional practices and activities that may not have been included in the sample. It is widely accepted that the main purpose of an evaluation study often defines and therefore limits its scope (Learning Technology Dissemination Initiative, 1998).

In order to avoid the mistakes of previous models, in particular, limiting the evaluation sample and using a grounded theory approach, Agostinho, et al., (2002) emphasised the importance of choosing an appropriate evaluation methodology (or several methodologies). To be truly generic however, it is argued that methodological considerations should incorporate a flexible system of evaluating different qualities and attributes similar to Marshall's (2007) model, in which each question is assessed in terms of the adequacy of its representation in a particular context. An extended review of evaluative methodologies is presented in the document titled "Review of Methodologies of Evaluating Technology-Enabled Learning and Teaching" (Quinton, Pachman, & He, 2010, Appendix 1).

As noted several times, students should also be considered valid stakeholders, and so one of the elements of the evaluative methodology should permit full expression of their concerns and encourage the right to influence all recommendations to adopt a particular educational technology. Alexander and Golja (2007) actively included students in a process of questionnaire refinement, but students were not considered decision-makers in the selection of different forms of technology. Instead the students' voices were used to map potential problems with existing systems and resources. A methodology that allows for automatic inclusion of faculty, students', IT experts' and administrators' voices holds genuine promise, not only for resolving potential methodological bias and problems, but also for instantiating a multiple-values evaluation approach as argued beforehand.

It is important to emphasise that the timing of evaluative activities (before, during, or after a certain technology is implemented) is also intricately related to the selected methodology and the format of the evaluation used. Owen (2006) recommends several methodologies and methods for the accumulation of evidence that may improve the final research findings, which may better inform subsequent iterative cycles.

Formative and summative aspects: establishing culture of evidence

From the outset of technology evaluation studies, a precedent was more or less established when it is considered that less than fifty percent of evaluative studies were based on formative activities (Alexander & Hedberg, 1994). As noted earlier, the rapid rate of change in technology often negates the relevance and meaning of applying timeframes to summative evaluation studies. An iterative formative evaluation process not only allows for a review of technological change, but also to establish a solid body of evidence from which to resolve the following question: What steps can institutions take to assist faculty members and other staff to gather and use evidence in order to improve learning and teaching practice?

The broader question of the scholarship of learning and teaching may involve the question of the number of participants / adopters as the more that become involved in enhancing their practice using technology, the greater an asset the TELT platform becomes for the university. As observed in education generally, the core conceptions of learning and teaching have changed "from imparting and acquiring 'content' to facilitating and constructing knowledge, a shift from 'product' to 'process'" (Cunningham et al., 1998, p. 32). Thus, an iterative formative approach may serve to address the evaluation requirements.

To be truly generic, an educational technology evaluation framework should also allow for several evaluative outcomes that include a capacity to reflect on the educational value and effectiveness of the framework itself. That is, in order that it can contribute to improving learning effectiveness, the TELT evaluation framework itself should also be evaluated to determine the degree to which it assisted to achieve such outcomes. Once again, an iterative evaluation process is required to ensure that specific activities may be modified in line with the results as they become available. Moreover, since all evaluation results will be disseminated, and the merit of the new framework is based in part on the evaluation outcomes, the framework requires summative evaluation strategies to allow for feedback and comment, and so ensure a balanced approach is consistently applied to the selection and use of educational technology.

Identifying the Major Components

Since the evaluation of technology-supported learning evolved at the same time as "the conflict between modern, positivist, experimental and postmodern, critical, case-based philosophies" occurred, many researchers regarded most technology evaluation studies at the time to be influenced by such philosophies and assumed to lack credibility as they were devoid of a valid systematic approach (Gunn, 1999, p. 186). Thus, given that the mandate for UNSW is to enhance learning through the effective use of technology, it was considered appropriate not to further such philosophical debates and instead ensure that all related policy, technical, and organisational factors

are guided primarily by pedagogy. In other words, all three evaluation components must directly reference the core attributes of learning and teaching scholarship as a primary focus.

A common problem encountered in the evaluation of educational technology is the absence of models that account for the learning process and the influences exerted by technology as students are learning (Atwell, 2006). As he observes, the European experience demonstrates that while the managerial (or business) component is usually well addressed in their models, little or no evidence exists of a dedicated focus on pedagogical factors. Again, from the European perspective, Attwell (2006) describes the use of management models for evaluation (a business component) and highlights the lack of integrative models that take into account learning processes and technology (iterations). The report confirms that educational policy and strategy are vital aspects of the overall evaluative process. These findings validate one of the concerns that initially led to the development of the TELT Evaluation Framework. These insights in turn add weight to the need to encapsulate all elements determined to be essential for effective educational technology evaluation.

The Curriculum Corporation (2004) framework incorporates three distinct layers of evaluation and takes into consideration business as well as technological processes. However, the pedagogical layer is relatively inadequate, not learner-centred, emphasises mainly the issues of security on the Web, and presents student engagement as a uni-dimensional process (yes or no). Moreover, the framework is aimed at evaluation of one particular technology with some attempts at comparative analysis. That is, it does not position a product within an array of technological tools that collectively can provide a greater variety of sophisticated functions for enhancing learning and teaching than can be obtained from a single multi-featured product such as a learning management system.

The broader issue of the scholarship of learning and teaching also involves the question of the number of participants / adopters based on the rationale that as more lecturing staff become interested in improving their teaching practice using technology, the greater the impact technology exerts on learning and teaching use and innovation across the university. Even in this regard the accepted definitions of learning and teaching are under question. As noted earlier, and acknowledged in education generally, the core conceptions of learning and teaching have changed “from imparting and acquiring ‘content’ to facilitating and constructing knowledge, in short, a shift from ‘product’ to ‘process’” (Cunningham, Ryan, Stedman, Tapsall, Bagdon, Flew, & Coaldrake, 1998, p. 32). In light of the importance of understanding the impact of continual change, it is further emphasised that an iterative, formative approach is appropriate to evaluating technology-supported learning and teaching.

One of the assumptions of evaluation framework designs that include a pedagogical component is that institutional policies and organisational structures are already in place for faculty and learners. Both the Boud and Prosser (2002) and the Dean, Biner, and Coenen (1996) studies rely on this perception. However, this position does not always hold true and therefore the business processes and administrative components must not be neglected in defining a generic evaluation model. Thus, the framework should also focus on the managerial (organisational), strategic goals of the university as well as its essential business processes that together underscore the overall effectiveness of the evaluation process. This factor is further highlighted by the realisation that the selection and use of information and communication technology (ICT) developments that support learning and teaching are no longer predominantly confined to isolated projects within academic departments and learning technology support units. Instead, ICT should be viewed as a core aspect of institutional policy (Conole, 2004).

The business processes component requires a comprehensive review of attributes such as meeting the demands for new job skills; cost control; the provision of widespread access to training/education opportunities; attracting/engaging new learners; and building an adequate labour force for community businesses. The business processes component also needs to include organisational and structural issues. In some cases, institutional strategy is “the sum of large numbers of independent actions by many faculty members and students across the college” (Ehrmann, 1995). The actual organisational structure or a planned structure can often turn out to be quite different from the projected structure,

and so if educational administrators did not plan for such inconsistencies, then the evaluation of administrative activities in relation to learning and teaching and faculty development can serve to identify crucial gaps.

One advantage of applying a generic evaluation framework is that it not only assists to improve organisational processes, but also resolve issues of organisational change that are often ignored in an additive approach to technologies evaluation. A detailed schema of business processes evaluation could reciprocate and put in place an explicit model for the management of change, relevant to tertiary education (Boys, 2002). Traditional evaluation models often do not explain why changes occur, how changes are different across settings, or how they relate to changes in the innovation (Bruce and Rubin, 1993).

The third evaluation component, technical, has been relatively well covered in research studies such as Obexer's (2005) model (Carmean & Brown, 2005) and the Curriculum Corporation model (Curriculum Corporation, 2004). Whilst covering technological factors, the Obexer model provides a narrow scope for evaluation in terms of including tools that are comparable to learning management systems, but the model does not allow for inclusion and subsequent iterative analysis of faculty specific tools and practices. Thus, another point to make that also relates to the educational / pedagogical component is the necessity to include "worldware" (Ehrmann, 1995) such as email, wikis, blogs, and authoring tools. All software is not designed for learning such as an LMS, but often serves an important function in the learning and teaching processes that should also be subject to a detailed evaluation. A generic framework should therefore aim to expand the technical component to include iterative technology-development and feedback cycles for a wide range of technologies whilst ensuring absolute adherence to the primary educational / pedagogical component.

Other useful findings include the need to consider the socio-cultural factors and governmental agendas that sit close behind the adoption and use of technologies (Conole, 2004); the concept of lifelong learning (Franklin, Yoakam, &Warren, 1996); and the characteristics of lecturers (Dean, Biner, & Coenen, 1996), that in part determine the key attributes that assist to devise the evaluation criteria for the pedagogical component.

As demonstrated by Britain and Liber (2004), and noted beforehand, an evaluation study should engage students in the process rather than viewing them as recipients of an educational service where all decisions are made on their behalf. The framework should view the learner as a self-regulated participant, responsible for his / her own learning as a collaborator and peer in an ongoing learning and evaluation process (Chickering & Ehrmann, 1996).

Integrative models and their shortcomings

Over time, there have been several large-scale attempts to create an integrative evaluation model for educational technology evaluation. The European project (Attwell, 2006) referred to earlier can be considered one significant example. This project however, presents an eclectic collection of evaluative frameworks that are applicable to the educational / geo-political contexts of various European countries. No generic or even comprehensive approach that is predominantly focused on learning and teaching, or technology enabled teaching and learning evaluation, can be identified in this collection.

One other attempt however, the Benvic Benchmarking system (2000, 2002) definitely extends beyond the narrow scope of a context-specific benchmarking methodology. It satisfies the conditions of a generic framework in presenting multiple pathways to technology evaluation, but recommendations made are based directly on the findings of previous field research and on the surface presents a sound theoretical base. Nevertheless, the model is too abstract for practical application. There are no concrete guidelines, or a variety of evaluation methods or ready-to-use checklists to be found in the Benvic deliverables document (Benvic Project, 2000).

The later report (Benvic Project, 2002) covers administrative/ infrastructure questions, but the pedagogical or technological variables are too broadly defined with little attention paid to providing details on the relevant pedagogies. A similar issue occurs with many large scale evaluative models that again give rise to questions that are too broad in their scope to resolve the development of useful pedagogical and technical criteria, while overemphasising the institutional structure factors (McNaught, Phillips, Rossiter, & Winn, 2000).

In summary, the alignment of a new evaluative framework to present and future learning and teaching needs requires that it draws on previous research in the field. As such, it has been firmly established that pedagogical, organisational (business processes), and technical requirements are viewed as core components in many evaluation studies (Conole, 2004). It is well established that from the commencement of technology evaluation studies, a precedent has more or less been established when it is considered that less than fifty percent of evaluative studies were based on formative activities (Alexander & Hedberg, 1994). As also noted, rapid technological change often undermines the validity of the resultant outcomes, particularly for summative studies. A formative evaluation process accommodates a review of technological change, and thereby provides a valid basis for determining the educational value of technology applications and the learning effectiveness of the online solutions and environments in which those technologies are applied.

A MULTI-DEFINITIONAL APPROACH TO EDUCATIONAL TECHNOLOGY EVALUATION

Pedagogical concerns over the actual learning effectiveness of various Educational Technology applications (often referred to as eLearning in literature) have increased alongside the meteoric rise in eLearning popularity and implementation. The chief criticism is that the increasingly feature-rich technological platforms are not properly developed for an appropriate learning context, often improperly integrated and under-utilised, and in many situations simply become an expensive means of indexing ordinary content, adding little educational value. Kiili's (2005) assessment is that "technologies, including computer games, are all too often used as substitute teachers that deliver information to learners rather than being used as learning tools to support an active learning process". Dublin (2003, p. 1) provides another dimension to these concerns:

> We keep talking about the technology, instructional design, and content, when the real issue is getting people to use what's developed, and getting organizations to truly integrate e-learning into everyday life and operations.

With rapidly changing information technology developments and the perennial need to restructure educational environments to adapt to an increasing demand for learning, comprehensive evaluation of educational technologies should be of primary concern for institutions attempting to embrace the new student-centred, Constructivist and social learning paradigms. Pohl's (2004) review of Baumgartner's (1995) work wisely reminds us that "every educational software is based on theoretical assumptions about the process of learning, whether explicitly or implicitly". These assumptions and their technological expressions can vary significantly and as a consequence yield quite different and often unexpected technology implementation results in different learning contexts. In essence, our need for evaluation can begin with the conclusions of Alexander (1999), writing on evaluating Higher Education Technology Innovations:

> "Without effective, scholarly evaluation, even well designed innovations are unlikely to achieve wider dissemination, and the potential benefits of [Communication and Information Technology] for learning in higher education are unlikely to be realised"

Not only do many educational technology theorists similarly suggest that comprehensive evaluation of eLearning systems is essential, but many such as Moore, Dickson-Deane, Galyen, Vo, and Charoentham (2008) strongly advocate that in the long run, "evaluators should have one tool, which can be manipulated to suit the type of e-Learning product." (Moore et al., 2008). Ideally then, such a tool should flexibly evolve with technological advancement. As will be explained, effective

information and communication technology (ICT) integration and unified evaluation in everyday teaching and learning needs to consider technology, content and pedagogy not in isolation, but more specifically as a three-dimension matrix of complex relationships that comprise a systemically structured teaching and learning whole..

THE EVIDENCED NEED FOR USABILITY, USEFULNESS AND EDUCATIONAL VALUE EVALUATION

In conjunction with the overarching aim of evaluating the actual pedagogical (learning) effectiveness of eLearning environments as assessed through various educational theories and methodologies, it is also important to consider usability and usefulness as two fundamental factors that assist to determine whether educational technologies are both conducive to effective learning and supportive of positive pedagogical strategies, and thereby possess educational value. Therefore, it is argued that prior to undertaking a comprehensive assessment of the learning effectiveness of an eLearning software application (or an online learning solution), a preliminary examination should also consider the usability, usefulness, and educational value as a function of educational technology use within the context of learning and teaching.

Throughout the world, usability has been increasingly recognised as a necessary element for moving towards the creation of useful and effective learning solutions for end-users. The "Ergonomics of Human System Interaction" framework (ISO9241), developed in 1998 by the International Organisation for Standardisation promoted the view that usability determines the extent to which "users of products are able to work effectively, efficiently and with satisfaction", therefore "the performance and satisfaction of the users provide direct measures of whether a product is usable in a particular context."

Specifically, from a teaching and learning perspective, Cheon and Grant (2009) explain: "Many educators develop online learning materials with little understanding of the importance of a functional, communicative, and aesthetically appropriate user interface (Metros & Hedberg, 2002). In fact, designing a usable and appealing interface is still challenging for instructional designers" (Chalmers, 2003). The following extracts from Cheon and Grant (2009) expand on this point (Cheon and Grant's references have been included in the Reference List for the reader's convenience marked '*'):

> The most important goal of designing a user interface is to reach learners more effectively, which has been studied in a number of ways. For example, capturing learners' attention and interest is one of the roles of a user interface (Hron, 1998; Parizotto-Ribeiro and Hammond, 2005; Szabo and Kanuka, 1998).
>
> Attractive displays can stimulate learners' engagement and improve learning performance (Metros and Hedberg, 2002). Also, an effective interface can facilitate navigation that contributes to transfer of information as well as communication between a user and a computer by representing and modelling a domain (Hron, 1998; Metros and Hedberg, 2002; Parizotto-Ribeiro and Hammond, 2005).
>
> In addition, the user interface scaffolds a student's ability to perceive, organize, integrate, and remember information (Chalmers, 2003; Haag and Snetsigner, 1993; Hannafin and Hooper, 1989; Szabo and Kanuka, 1998). While good interface design can facilitate effective learning, poor design could hinder learning. For example, poorly organized and designed interfaces could inhibit formation of schema and contribute to disorientation (Chalmers, 2003; Cheon and Grant, 2009).

In short, if the usability of a particular technology application or learning solution is a hindrance to users' learning and teaching experiences and if not evaluated separately, the entire evaluation process will present a biased impression of the overall educational *value* and as a consequence the learning effectiveness of the evaluated product.

Cheon and Grant (2009) continue on to provide an example: "an inconsistent and poorly designed interface may produce confusion that causes learners to overly attend to the organisation and navigation of the instructional unit." Benson, Elliot, Grant, Holschuh, Kim, and Lauber's work (2002) (2002) similarly point out that "usability is critical to a learner's experience with an e-learning program, and the effectiveness of even the most instructionally sound programs will be decreased if the learner's experience suffers from problems related to navigation, orientation, visual appeal, and other usability criteria" (Reeves & Carter, 2001). Indeed, the **usability of online learning systems has a significant documented positive correlation (r=0.83) with student learning outcomes (Meiselwitz & Sadera, 2006), and thus its educational value and subsequent learning effectiveness.**

Related educational technology research conducted by Crowther, Keller, and Waddoups (2004) notes from the work of Kim, Douglas, Orkand, & Astion, (2001) that "if usability problems impair student performance, it is difficult to assess the true educational value of the application." Crowther hypothesises not simply a generic evaluation framework where assessment ranges from basic interface usability to pedagogical integration, but one that also supports distinct layers of iterative testing: "Poor student performance demonstrates that the product is not achieving the design team's objectives, but it does not identify whether the fault lies with the educational concept, the presentation, the interface or all three. Iterative user testing can distinguish between conceptual, pedagogical and interface problems, thus preventing poor design from skewing evaluation measurements of educational effectiveness."

DEVELOPING A GENERIC USABILITY, USEFULNESS, AND EDUCATIONAL VALUE FRAMEWORK

Across the entire range of application-specific usability, usefulness and educational value (UUE) evaluation methodologies reviewed in this study, evidence emerged of similar collections of usability attributes that are of specific importance to users' ability to successfully access and use educational technology applications and online learning solutions in the manner intended by the lecturer / facilitator.

By narrowly focusing only on individual applications, the literature often confuses what are essentially usability evaluations of functions and features specific to that application and then extrapolate the findings to form a general '*usability*' evaluation without taking into account the full range of factors that otherwise may lead to entirely different conclusions. Although such ad-hoc usability characteristics can be extracted to inform a generic evaluation study, there are other important factors such as usefulness, and educational value that when combined provide a more detailed insight across a wide range of applications and is thereby more relevant to a wider audience with its familiar terminology and scope.

In developing an educationally effective generic UUE framework suitable for a wide variety of technology applications, learning solutions, and users, it is possible to begin by reviewing abstract usability theory, with exemplary extractions from specific evaluation research such as the broad-scoping, traditional usability heuristics pioneered by Dr. Jakob Nielsen (Nielsen, 1993; Nielsen, 1994), and outlined in Table 1. It is argued that heuristic sets such as Nielsen (1993) and Nielsen (1994) have withstood the test of time, having been widely adapted and applied to numerous applications both within the educational technology / eLearning context and to more general software and hardware usability evaluations.

Table 1- **Nielsen's Five Usability Attributes (1993) and Ten Usability Heuristics (1994)**
(Nielsen, 1993; Nielsen, 1994)

Learnability	The system should be easy to learn so that the user can rapidly start getting some work done with the system
Efficiency	The system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible

Memorability	The system should be easy to remember, so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again
Errors	The system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Further, catastrophic errors must not occur
Satisfaction	The system should be pleasant to use, so that users are subjectively satisfied when using it; they like it

1	Visibility of System Status
2	Match between system and real world
3	User control and freedom
4	Consistency and Standards
5	Error Prevention
6	Recognition rather than recall
7	Flexibility and Efficiency of Use
8	Aesthetic and minimalist design
9	Help Users recognise, diagnose, and recover from errors
10	Help and Documentation

The assessment of usability through heuristics (such as Nielsen's for example) allows for the specific evaluation of individual educational technologies without the need to administer questions to staff and students and thus affords more effective use of time and resources by focusing only on their unique characteristics and attributes. By not burdening the student and staff respondents twice with lengthy and 'obvious' generic UUE questioning at the application-specific evaluation stage, it is possible to obtain higher quality data. It is at the later stages of the evaluation cycle, when focusing on the application of an educational technology tool within a pedagogical context, a course, and an overall learning environment that context-specific evaluations by students, learning facilitators, and course administrators can best determine the real learning effectiveness gains from the context-specific UUE elements, and then compare and contrast the results with the assessment that was made of the generic 'look and feel'.

The use of variants of Nielsen's heuristics to evaluate the usability and usefulness of an educational technology application, both general and context-specific, can be found as recently as 2009 in Sim, Read and Cockton's (2009) work on CAA (Computer Assisted Assessment) evaluation in which they concluded that the effectiveness of Nielsen's heuristics is worth evaluating in the context of CAA. For this study, the decision was made to use Nielsen's heuristics as they are the most generic and widely applied (Sim et al., 2009).

In their comprehensive work exploring usability and human-computer interaction Shneiderman and Plaisant (2004) recommend that designers and evaluators focus on exploring system interaction through five measurable usability dimensions (see Table 2), which in turn support two of the three ISO9241 ergonomics/usability goals of *efficiency* and *satisfaction*.

Table 2 - **Key Usability Measures** *(as per Shneiderman & Plaisant, 2004)*

Time To Learn	How long does it take for typical members of the user community to learn how to use the actions relevant to a set of tasks?
Speed of Performance	How long does it take to carry out the benchmark tasks?

Rate of Errors by Users	How many and what kinds of errors do people make in carrying out the benchmark tasks? Although time to make and correct errors might be incorporated into the speed of performance, error handling is such a critical component of interface usage that it deserves extensive study.
Retention Over Time	How well do users maintain their knowledge after an hour, a day, or a week? Retention may be linked closely to time to learn, and frequency of use play an important role.
Subjective Satisfaction	How much did users like using various aspects of the interface? The answer can be ascertained by interview or by written surveys that include satisfaction scales and space for free-form comments.

A comparison of Nielsen's (1994) ten usability heuristics with Shneiderman and Plaisant's (2004) usability dimensions indicates common measures that can be adapted to evaluating generalised educational technology applications and environments, and reorganised as theoretical stages within a **Usability, Usefulness and Educational Value (UUE) heuristic life-cycle**. Questions derived from this heuristic life-cycle can then be used to easily evaluate a wide gamut of user interaction outcomes across the entire time-span of an application's or a learning solution's use.

Again and again, it is clear that a similar set of usability elements is prioritised not only in abstract theory, but also in industry standard documentation from leading technical laboratories such as Sun Microsystems, which concluded that usability is comprised of six general attributes: *utility*, *learnability*, *efficiency*, *retainability*, *errors* and *satisfaction* (Sun Microsystems, 2001). Accordingly, all these recurring theoretical themes are thoroughly assessed across the five-stage life-cycle.

As another example, consider that Cheon and Grant (2009) identified "three considerations for developing [online] user interface: too much information on a screen (Lyons, 2001; Khentout, Harous, Douidi & Djoudi, 2006), the impatience of Web users (Lyons, 2001) and a tendency to become lost and disoriented (Lyons, 2001; Khentout et al., 2006)." (Cheon and Grant's references have been included in the Reference List for the reader's convenience marked "*"). Here, the Generic UUE Evaluation would need to incorporate Cheon and Grant's concerns as discrete points of interest under thematic blocks such as "minimalist interface"/ "aesthetics", "speed of user task performance / application responsiveness" and "logical navigation" respectively. Whilst many such theoretical recommendations (such as Cheon & Grant, 2009) exist, they are rarely collated together into a generic framework, nor do they guide the evaluation into effective pathways for further action. It would thus add value to the existing literature to design a generic UUE framework that is both easy to administer and can effectively generate valuable and actionable outcomes.

Over the course of the last two decades of living in an increasingly technologically-enabled world, it is likely that many users have become accustomed to certain instances of 'standard' software and hardware that behave and look in similar and expected ways, heuristic theory considerations notwithstanding. The general layout of the Microsoft Windows operating system is a familiar example. This is something that a Generic UUE Evaluation should also take note of: by monitoring changes in user interface trends set by established hardware and software providers, the inherent iterative nature of the framework can allow it to develop alongside the prevailing trends. Therefore, a Generic UUE Framework should appropriately derive and evaluate for certain usability criteria that have, by evolutionary processes, become the modern status quo for user interface satisfaction with the lowest learning curve. Following on from this example, a model could assess subconscious considerations such as the position of a button to close a window – now naturally top left or top right in software applications. Any alternative positioning of such a button, even if clearly labelled, could thus distress a user intuitively scanning the top corners of an application window.

Pathways have also been established into adapting the Nielsen (1994) heuristics towards various assessments of the usefulness and educational value of generalised eLearning environments. Reeves and Hedberg (2008) build on Benson et al's., (2002) scale, for example. Benson et al., (2002), after finding that "Nielsen's original ten usability heuristics were insufficient for e-learning programs",

recommended adding seven extra heuristics that are "designed to be more closely focused on e-learning programs." Reeves and Hedberg (2008) went on to add an extra three – Performance Support Tools, Learning Management and Content (shaded grey in Table 3) – for a total of ten.

Table 3 **- Ten Additional eLearning-specific Heuristics and Suggested Sample Questions**
(as per Reeves & Hedberg, 2008)

Interactivity	The eLearning program provides content-related interactions and tasks that support meaningful learning.
e.g.	a. Does the e-learning program provide too many long sections of text to read without meaningful interactions? b. Does the e-learning engage the learner in content-specific tasks to complete and problems to solve that take advantage of the state-of-the-art of e-learning design? c. Does the e-learning program provide a level of experiential learning congruent with the content and capabilities of the target audience?
Message Design	The eLearning program presents information in accord with sound information-processing principles.
e.g.	a. Is the most important information on the screen placed in the areas most likely to attract the learner's attention? b. Does the e-learning program follow good information presentation guidelines with respect to organisation and layout? c. Are graphics in the e-learning program used to clarify content, motivate, or serve other pedagogical goals?
Learning Design	The interactions in the e-learning program have been designed in accord with sound principles of learning theory.
e.g.	a. Does the e-learning program provide for instructional interactions that reflect sound learning theory? b. Does the e-learning program engage learners in tasks that are closely aligned with the learning goals and objectives? c. Does the e-learning program inform learners of the objectives of the program and remind them of prior learning?
Assessment	The e-learning program provides assessment opportunities that are aligned with the program objectives and content.
e.g.	a. Does the e-learning program provide opportunities for self-assessments that advance learner achievement? b. If appropriate to the context, do assessments provide sufficient feedback to the learner to provide remedial directions? c. Are higher order assessments (e.g., analysis, synthesis, and evaluation) provided wherever appropriate rather than lower order assessments (e.g., recall and recognition)?
Media Integration	The inclusion of media in the e-learning program serves clear pedagogical and/or motivational purposes.
e.g.	a. Is media included that is obviously superfluous, i.e., lacking a strong connection to the objectives and design of the program? b. Is the most appropriate media selected to match message design guidelines or to support specific instructional design principles? c. If appropriate to the context, are various forms of media included for remediation and/or enrichment?
Resources	The eLearning program provides access to all the resources necessary to support effective learning.
e.g.	a. Does the e-learning program provide access to a range of resources (e.g., examples or real data archives) appropriate to the learning context? b. If the e-learning program includes links to external World Wide Web or Intranet resources, are the links kept up-to-date? c. Are resources provided in a manner that replicates as closely as possible their availability

	and use in the real world?
Performance Support Tools	The e-learning program provides access to performance support tools that are relevant to the content and objectives.
e.g.	a. Are performance support tools provided that mimic their access in the real world? b. Provided the context is appropriate, does the e-learning program provide sufficient search capabilities? c. Provided the context is appropriate, does the e-learning program provide access to peers, experts, instructors, and other human resources?
Learning Management	The e-learning program enables learners to monitor their progress through the material.
e.g.	a. Does the learner know what he/she is doing and how he/she is doing within various parts of the e-learning program? b. Does the learner perceive options for additional guidance, instruction, or other forms of assistance when it is needed? c. Does the learner possess an adequate understanding of what he/she has completed and what remains to be done within any specific unit (e.g., a course) of e-learning?
Feedback	The eLearning program provides feedback that is contextual and relevant to the problem or task in which the learner is engaged.
e.g.	a. Is the feedback given at any specific time tailored to the content being studied, problem being solved, or task being completed by the learner? b. Does feedback provide the learner with information concerning his/her current level of achievement within the program? c. Does the e-learning program provide learners with opportunities to access extended feedback from instructors, experts, peers, or others through e-mail or other Internet communications?
Content	The content of the e-learning program is organised in a manner that is clear to the learner.
e.g.	a. Is the content organised in manageable modules or other types of units? b. Is the content broken to appropriate chunks so that learners can process them without too much cognitive load? c. Does the e-learning program provides advanced organisers, summaries, and other components that foster more efficient and effective learning?

In light of the above, to evaluate the usefulness and educational value of an application with consideration to its usability, Reeves and Hedberg's (2008) additional educational technology-specific heuristics can also be integrated into the UUE life-cycle. Consequently, it is possible not only to evaluate the usability of distinct elements of the application across the time span of its natural use (importantly including elements most specific to educational technologies), but also to assess **the usefulness and educational value that is produced through their effective, combined usage within a learning and teaching context.**

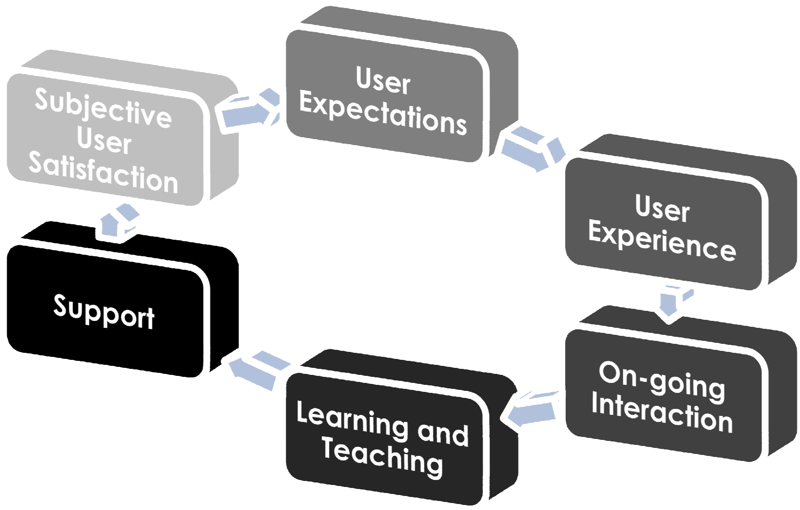
In another example of what occurs when usability is assessed without reference to usefulness and educational value, Sim, Read, and Holifield, (2007) assessed undergraduate students' (*N=44*) satisfaction with three different Computer Assisted Assessment (CAA) applications across usability attributes such as "accessing and finishing test", "navigation within the test", "visual layout" and "interface for answering questions". They concluded that "there does not appear to be a single attribute that influences students' preference for a particular CAA environment" (Sim et al., 2007) – although, paradoxically, every user may still have a discrete preference, and consequently a varied learning experience.

Therefore, on a practical level, the **six-stage heuristic life-cycle** illustrated in Figure 1 forms a high-level, evidence-based starting point that informs the creation of tailored and thematically grouped questions that revolve around the entire range of technology application usability experience factors, as well as the creation of the survey instruments that drive the validation processes which underpin

the effectiveness of a generic evaluation framework. The questions can be designed to assess the most common usability, usefulness, and educational value attributes, and typical usability problems as identified through the extensive evaluation literature review undertaken in 2010 and 2011, which builds on the review undertaken in the supporting document titled "Evaluation of the TELT platform: Essential elements and methodologies" (Quinton, Pachman, & He, 2010).

The crux of a Generic Usability, Usefulness and Educational Value Evaluation framework therefore becomes an assessment of the **nexus of usability, usefulness, and educational value dimensions as an evaluation across the various ad-hoc elements that arise during the natural use of learning and technology applications over time.** It is this nexus that creates the ephemeral, intuitive 'look and feel' of user experience and satisfaction that has been so inherently difficult to individually isolate and evaluate, and also drives forward the perceived usefulness and educational value of an application in the mind of a staff or student respondent. Assessments of integrated learning effectiveness and specific learning and teaching environments however, are best addressed at a later stage through a comprehensive context-specific evaluation of a technology, as guided by the theoretical sub-layers outlined in Tables 5, 6, and 7.

Figure 1 - **Generic UUE Evaluation – The Six-Stage Life-Cycle**



DESIGNING A GENERIC UUE EVALUATION METHODOLOGY

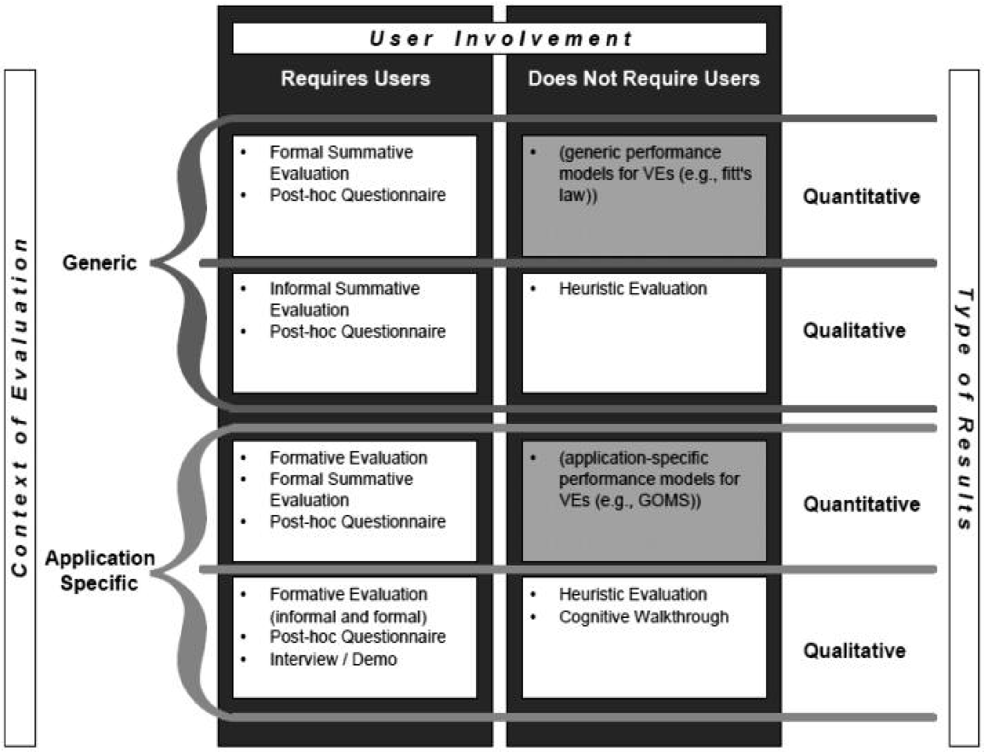
Based on Bowman, Gabbard, and Hix's (2002) methodology selection model (see Figure 2) and the eventual evaluation output requirements, the entire Generic UUE Evaluation across the usability stages can be conducted as a <u>standardised quantitative questionnaire</u> for staff and student respondents to undertake as the first part of any user-centric evaluation instrument set, prior to any application specific or context specific inquiry – which, judging by the evidence, are generally more information-rich and open-ended in their output. A generic questionnaire can be useful to familiarise the user with the evaluation process and provide an easy, relevant and fast entry point for feedback collection and further evaluative research. Ideally, the structured aggregate responses will provide valuable data unbiased by application specifics, whilst the questions should be familiar to many users of varied technologies, by using recognisable jargon, descriptions and concepts.

By subsequently leveraging the Generic UUE quantitative data from an appropriate number of staff and student respondents, valuable insights can be gained. Most importantly, systemic usability issues that have not been picked up by the vendor/developer, or at other stages (Stages 1 and 2 – see Figure 4), can be readily and most cost-effectively identified (Hom, 1998), and the results compared across a wide range of other technological applications. Similarly, having performed a generic evaluation

across a range of applications, the data may reveal that systemic usability issues could also for example, be a function of user technological training. A Generic UUE Evaluation with end-users can also be performed at various timepoints – perhaps before and after student training or staff development, or a university preparation programme – to explore the extent to which training and education plays in the user's perception of and interactions with an application, and the subsequently derived educational value factor.

It is also important to tailor the question coding schema of the Generic UUE Evaluation to explore the aims of the overall framework. Provision should exist for a range of response scales, such as: the useful Likert scale similar to Nielsen's (1994b), visual identification of usability problems, and domain-specific coding based on the recent reasoning of Sim et al., (2009) – a personal satisfaction / perceived personal educational value mini-scale for usability within the context of educational value.

Figure 2 - **A Classification of Usability Evaluation Methods for Virtual Environments**
(as per Bowman et al.,2002)



THE STAGES AND SUB-LAYERS OF THE TELT EVALUATION MODEL

The Three Stages

The TELT Evaluation framework incorporates three distinct, yet highly interrelated (primary) stages (see Figure 4). Stage 1 focuses on vendor assessment; Stage 2 involves a pre-pilot scoresheet test; and Stage 3 initiates an ongoing educational evaluative cycle, first to ensure the targeted technology application (or solution/s) complies with the established criteria (for subsequent inclusion or exclusion from the TELT platform) and second, to evaluate the learning effectiveness of the approved educational technology to align with the need to enhance and improve online learning and teaching practices at UNSW. The Stage 3 cycle also feeds into a broad-level cycle of evaluation of the entire TELT Evaluation framework to ensure continual refinement and modification of all three stages in

terms of their respective criteria, their specific interrelationships, and their interdependencies with learning and teaching.

In Stage 1, prospective vendors or open source providers are requested to produce evidence of their long-term viability, business processes, and commitment to the advancement of the product. Vendors are assessed using a multi-stepped screening approach. Subsequent potential implementation of the product through the TELT platform is conditional upon successful completion of the pre-pilot (Stage 2) and full pilot (Stage 3) evaluations. The rationale behind the Stage 1 evaluation is to optimise the university's expenditure and time commitments in determining the suitability of a vendor's product prior to commencement of the more detailed Stage 2 and Stage 3 evaluations (see Figure 3).

Stage 2 involves a 'pre-pilot' evaluation of a software solution. The process is informed by the outcome of Stage 1 in terms of determining the types and comprehensiveness of the criteria to be applied. A pre-pilot scoresheet (producing a quantitative output) has been designed to assess and compare products (if more than one shortlisted). All products are evaluated against the existing TELT Platform educational needs and planned directions with respect to the range of features offered (functional), the implications for effective integration with UNSW's business processes, and the product's technical specifications for a detailed comparison with the university's IT infrastructure requirements. If the product (or products) scores highly, it is selected for the full pilot evaluation in Stage 3.

The rationale behind a pre-pilot evaluation is the rapid rate of technological change. As noted, the situation that occurs when "once something is ...evaluated... it is already obsolete" (Bruce, 1999) began to devalue the integrity of educational technology evaluation from the outset of the 1990s. The results of longitudinal studies in this area often became redundant by the time a technology solution is ready for implementation, again due to the rapidity of technological advancements and innovation (Baker and Herman, 2003). The short period of time (2 to 4 weeks depending on the size and complexity) required to complete the evaluation scoresheet reflects how Stage 2 is designed to mitigate these problems.

Stage 3 determines the suitability of online learning products and software solutions for inclusion in the TELT platform and once approved, the subsequent release to the university community. It introduces a three (sub)-layered, full pilot evaluation that requires several evaluative aspects to be incorporated into a unified process. On one level, it observes a systematic approach to a detailed evaluation of three crucial elements: the pedagogical imperatives, the institutional, faculty, and school level business processes, and the technical requirements of the TELT platform. On a much broader and deeper level, the framework affords genuine systemic flexibility in selecting the types and complexity of criteria, and the analysis methodologies / approaches that are applied to each of the preceding three elements. The evaluation processes outlined here parallels the approach embraced by Marshall (2007) in the construction of his eLearning Maturity Model (eMM). In describing the "Dimensions of capability" of the eMM model, Marshall (2007, p 6) writes:

> A key development that arose from the evaluation of the first version of the eMM is that the concept of levels used was unhelpful (Marshall and Mitchell, 2006). The use of levels implies a hierarchical model where capability is assessed and built in a layered way. The key idea underlying the dimension concept in contrast, is holistic capability. Rather than the model measuring progressive levels, it describes the capability of a process from synergistic perspectives. An organization that has developed capability on all dimensions for all processes will be more capable than one that has not. Capability at the higher dimensions that is not supported by capability at the lower dimensions will not deliver the desired outcomes; capability at the lower dimensions that is not supported by capability in the higher dimensions will be ad-hoc, unsustainable and unresponsive to changing organizational and learner needs.

As indicated, the evaluation criteria applied in Stage 3 (also referred to as sub-layer 1) consist of three interrelated layers: Pedagogy (L&T), Business, and Technical. The layers (illustrated in Figure 4) collectively serve to establish a multi-referenced foundation on which to systematically evaluate targeted technology applications and solutions. The Pedagogy Layer is emphasised first and foremost as it directly supports the visions and principles drawn from the UNSW Graduate Attributes, Learning and Teaching Principles and the 16 Guidelines for Learning and Teaching Practice. To ensure full representation of all UNSW L&T related perspectives and needs, the pedagogical evaluation criteria used in this layer is further informed by the preferred teaching approaches and models applied by faculties and schools.

The business layer is primarily concerned with the range of linked activities that create benefit and value for users and the institution through processes and strategies that transform learning and activities (input) into more effective and valuable outcomes (output) in the context of deploying eLearning applications. Constant technological change and refinements to learning and teaching approaches will always mandate change in business processes. Thus, there is a need not only to evaluate TELT applications in line with the agreed business processes identified for the TELT Platform, but also to subject the business processes per se to an iterative evaluation strategy to determine whether further refinement is required.

The Stage 3 Business Layer seeks to create value for users and the institution by integrating teaching and learning innovation with organisational structures and procedures, and reviewing business processes and operations, to create valuable and effective strategically aligned TELT outcomes. Amongst other functions, it:

- identifies formal business processes and procedural requirements, standards and practices for organisational change across both the University-wide and Faculty specific contexts
- focuses on and assesses the business and organisational policies, guidelines, reports, staff and their responsibilities, procedures of relevant services and operation, planning of future operation, quality control, training and support, and vendor relationships and management as relevant to their own iterative evaluation and refinement, and their relationship with the TELT Evaluation Framework and teaching and learning practices, and
- informs the evaluation of any potential business processes and organisational change procedures, issues and pitfalls, and operational best practices to deliver successful, and institutionally compliant, teaching and learning improvements and innovative outcomes.

Thus, the Business Layer aims to deliver a framework for on-going assessment of and compliance with business and procedural requirements, operational feasibility, and the effective relationship between eLearning implementations and organisational standards and institutional practices. It recognises the formalised processes of organisational procedures and change at UNSW that are informed first by pedagogical and then by technological factors. Any new propositions related to institutional business needs and procedural changes that emerge during the evaluation cycle further inform the refinement of the business and technical layer criteria as determined to be appropriate.

In general, the TELT Platform business processes evaluation cycle will need to focus on: policies, guidelines, reports, staff and their responsibilities, procedures of relevant services and operation (for example, staff and student training, support, educational design), planning of future operation, quality control, vendor relationship and management.

The technical layer of the TELT Evaluation Framework applies to the technical standards as specified and observed by UNSW, and the related issues and needs that require resolution in relation to the technical attributes of the TELT Platform. Where the technical standards are concerned, there is a need to identify the industry standards that are common to the eLearning software applications approved for the TELT Platform as applied to the various categories of applications outlined in this document. These categories will change and expand as new technical standards, requirements and

applications emerge. Therefore, the technical requirements for TELT Platform applications must also be subject an iterative evaluation process.

In essence, the Stage 3 Technical Layer explores information technology within the University context. Amongst other functions, it:

- identifies technical eLearning standards and industry platforms
- assesses the technical needs, capabilities, strengths and weaknesses of institutional users and IT infrastructure, and
- informs the evaluation of any technical eLearning integration potential issues and pitfalls, and also information technology best practices to deliver successful platforms conducive to achieving effective learning and teaching processes.

Thus, the Technical Layer aims to deliver a framework for on-going assessments of eLearning technical compliance, feasibility, and effective integration with the whole community of users and institutional information technology.

Full compliance with technical standards and ongoing resolution of the issues and needs that arise in the use of the application features ultimately serves to ensure seamless interoperation, integration, adaptation, compatibility, and security. Moreover, rigorous observance of the technical compliance process also ensures thorough insight and familiarisation of each product's features may be gained as a result of using and testing to determine their strengths and weaknesses.

As for the technical issues, there is a need to identify the problems that commonly occur in the use of different types of eLearning products and to evaluate how these products are designed to minimise any negative impacts. Technical issues are not just about the applications. They also relate to any problems that are associated with the implementation process that includes for example, the user's required technical knowledge, the internal technical support that is needed, and considerations of users that require access to accessibility options.

Finally, the Technical Layer addresses the need to account for the technical requirements, standards and protocols established at UNSW that are informed first by the pedagogical and then the business process criteria. Any new propositions related to technological advancements and infrastructure changes that emerge during the evaluation cycle inform the refinement of the technical and business layer criteria, again as determined appropriate. For both the Business and Technical layers an extensive set of evaluation criteria have been established for completion in collaboration with the relevant departments within UNSW.

As made explicit earlier, the rational for adopting a three layered approach is that the pedagogical, technical, and organisational aspects represent the fundamental sets of factors required to ascertain the educational suitability of any learning technology (Conole, 2004). As the evaluation framework evolves through additional iterations of refinement and modification, it is likely that more layers may be required. For the present, the three layers outlined here provide an absolute minimal, albeit substantial basis on which the TELT evaluation framework is established.

The Three Sub-Layers

Embedded into the Stage 3 (Pedagogy Layer) is an iterative (three) sub-layered evaluation process to ensure continual modification and refinement of all criteria used in each of the three primary layers (see Figure 5). A closer examination of the Pedagogy layer illustrates the three learning and teaching related evaluation sub-layers. Once a technology has been approved for inclusion in the TELT platform, it is at this point that the systemic sub-layer cycle is triggered. The essential function of each sub-layer is to evaluate the:

1. educational value, usability, and usefulness of a technology application for inclusion in the TELT platform (sub-layer 1)
2. learning effectiveness of approved TELT applications (sub-layer 2), and
3. learning environments that are supported by approved TELT applications (sub-layer 3).

Sub-layer 1 evaluates the educational usefulness and value (the usability) of a technology application, tool, or solution (juxtaposed with the Stage 3 business and technical layers). The purpose of sub-layer 1 is to avoid the risk of the widespread "technology driven pedagogy" mistake. Often faculty are advised to move "beyond intuitive value" and "adapt to the changing landscape" (Carmean and Brown, 2005, p. 5) in the implementation of new technologies, while the reasons behind the change such as technological progress or the commercial interests of LMS developers are not made clear. If the technology application addresses the criteria for the three primary layers, then it will be assessed and approved for release to staff through the TELT Platform.

The premise on which sub-layer 1 rests is that technology is not merely an inert tool - it possesses inherent values (either in how it is designed or the way it functions) that may directly or indirectly influence pedagogical affordances. In other words, technology exerts its own pedagogical orientations (Britain and Liber, 1999). With this factor in mind, the assessment of educational value, usability, and usefulness of an educational technology is crucial to determining whether it is suitable for inclusion in the TELT Platform.

Sub-layer 2 evaluates the learning effectiveness of an approved TELT application. This sub-layer is intricately related to the educational values of the instructor, school, faculty; learners' characteristics; and, the learning environment. As such, sub-layer 2 is a crucial component of the iterative sub-layer evaluation cycle in that it is dependent on and also informs the evaluation methodologies and relevance of the criteria used in sub-layers 1 and 3. Highly graphic, visual snapshots of this type of evaluation can be found in the examples presented by Reeves and Laffey (1999). Although the authors have limited the scope of their evaluation to problem-based engineering courses, their evaluative approach is readily transferable to pedagogies and learning environments developed for other disciplines.

Sub-layer 3 evaluates the learning environments (the online course design) in which TELT applications are employed. It incorporates many of the best practices and recommendations advanced by the educational technology field over the last decade (see Graham, Cagiltay, Lim, Craner, & Duffy, 2001, for an example). These approaches may be blended with or added to existing faculties' and schools' preferred design and delivery models, and practices. In this way, the flexibility in methodology choice afforded by sub-layer 3 serves to enhance the learning effectiveness of online environments as well as inform the evaluative methodologies and relevance of the criteria applied in sub-layers 1 and 2 as part of the ongoing cycle of refinement and modification. Table 6 summarises the strategic outcomes to be derived from the sub-layer 1 to 3 evaluation sub-cycle.

THE SURVEY INSTRUMENT: SUB-LAYER 1

Given that the overall goal of Stage 3 is to determine the educational value and learning effectiveness of a given technology, it was decided that the survey instrument should be applicable to a broad range of educational technologies particularly where a comparison of different technologies (as opposed to the evaluation of a single technology) is required. This aspect addresses the need for valid and reliable measurement tools for evaluating all educational technologies. Finally, the instrument allows the synthesis of the contextual factors (mentioned below) in order to provide a clearer picture of the student population in relation to their perceptions of educational technologies as well as their overall ideas about learning and teaching.

Sub-layer 1, Stage 3 of the TELT Evaluation Framework is aimed at assessing the effectiveness of a given technology for pedagogical use, by way of exploring its usability, usefulness and educational value. The evaluation procedure involves the use of a survey instrument that is designed to reflect

staff and students’ opinions about educational technologies in a holistic context (that is, accounting for their previous experience with technology; preconceived notions of learning; resistance to implied authorities in relation to learning and teaching; preconceived notions about eLearning; and the flexibility of the technology application in relation to learning and teaching). The implementation of a staff and student survey in this sub-layer provides direction for making informed recommendations on perceived trends in technology adoption and for tailoring solutions to accommodate staff and student needs.

The instrument permits the synthesis of contextual factors such as students’ technological background and their understanding of teaching and learning processes, with the identified benefits and enhancements that are offered by a given technology. In turn, this synthesis provides an insight into students’ perceptions of educational technologies and their overall ideas about learning and teaching along with the usability, flexibility, and emotional affordances of a chosen technology as applied to students’ actual attributes and preferences.

The evaluative process is also designed to include an iterative component and represents the first iteration of the survey instrument started this process. The 2nd iteration is aimed at further refining the existing subscale and creating linkages between sub-layer 1 and sub-layers 2 and 3 by introducing relational elements, which are termed subscales for the purposes of this study (see Table 5).

THE SCALE

Nielsen’s heuristics (1994) and Benson et al’s., (2002) usability and instructional design heuristics served as the basis for the development of the usability survey items that are incorporated into the TELT Evaluation framework. As noted, Nielsen’s heuristics (1994) have been widely employed in usability evaluation, including the evaluation of educational applications. For example, Peng, Ramaiach and Foo (2004) performed a Nielsen’s heuristic-based user interface evaluation (N=88) of the electronic media services gateway at Nanyang Technological University (Singapore). Sim et al., (2009) also adapted Nielsen’s model (1994) as a framework for “synthesising a set of domain specific heuristics for evaluating CAA applications” (Sim, et al., 2009, p. 204). For the assessment of the emotional aspects of interaction with the system, a combination of Benson et al., (2002); Saadé, He & Kira (2007); and Venkatesh, Morris, Hall, Davis, Davis & Walton’s (2003) developments were applied. Benson et al., (2002) also used an expert panel comprised of eight members to evaluate and apply usability heuristics to learning software.

Saadé et al., (2007) studied the dimensions of online learning and system usage through the four constructs of student attitude, affect, motivation and perception with undergraduate students (N=105). They concluded that the affect and perception dimensions in their model have strong measurement capabilities. In addition therefore, four items were designed to assess student resistance to implied authorities in relation to one’s teaching and learning on a 7-point Likert response scale (from 1 = strongly disagree to 7 = strongly agree). The inclusion of the preconceived notions about eLearning items were influenced by the work of Ainley, Banks and Fleming (2002) and Baylor and Ritchie’s (2002) educational computer usage categories. Finally, two other items were designed to assess flexibility of application in relation to teaching and learning. Two content experts were then asked to review the scale separately. Based on their feedback, the scale was revised and finalised.

In completing the survey, participants were required to express their degree of agreement or disagreement with each statement by selecting one of the following seven options: 1 – Strongly Disagree; 2 – Disagree; 3 – Somewhat Disagree; 4 – Neither Agree or Disagree; 5 – Somewhat Agree; 6 – Agree; and 7 – Strongly Agree. Participants were also invited to provide non-mandatory short qualitative feedback alongside any particular statement. Again, a seven-point Likert scale was chosen as the most appropriate response tool. Many respondents were deemed to be familiar with the Likert scale, which was selected for its traditional usefulness, clarity of understanding, and statistical robustness. Measurement studies are quite unanimous in claiming the seven-point scale as an optimal

scale in relation to the number of response options, and as the scale with maximum reliability (Cox, 1980; Preston & Colman, 2000).

One of the practical considerations made during the initial planning stage included the size of the survey: it was limited to approximately thirty items given that participation was voluntary and the majority of students were not receiving additional course credits for participating in the evaluative activities.

A total of 152 items was originally developed for survey purposes. However, as the agreed limitation was confined to around 30 questions, the total number of questions related to the contextual factors was assumed to be around 22 items, and 8 demographic items (learners' previous experience with technology) served as a basis for a division on learners groups (see TELT educational technologies pilots: brief summary of findings doc). Refer to Table 4 below for a full list of the survey questions in use for Stage 3, sub-layer 1.

Table 4 – Sub-layer 1 Survey Questions

Item No.	Question
1.	Technology enabled learning (or teaching) is attractive to me in my daily life
2.	I integrate information technology into various aspects of my day to day life
3.	I always try new information technology tools to find innovative solutions to my learning or teaching activities
4.	As I use an information technology application I tend to think about the different ways in which I could personally apply it to my learning or teaching, beyond its promoted functionality
5.	I tend to consciously think about how I can use information technology to improve my learning or teaching approach and learning or teaching processes
6.	I am the type of person who would rather have a strong opinion than no opinion at all
7.	The way the information and resources are organised within the application easily integrates with the way I naturally organise or access my information and resources
8.	The process of searching for, finding and accessing the information I needed was clear and intuitive
9.	The application allows me to perform my tasks quickly
10.	In considering the usability and usefulness of the features and functions of this application, how strongly would you agree that you thought there were NO problems that you particularly identified
11.	I think the application has a pleasant, appealing look and feel
12.	The application integrates well into my learning or teaching approach and processes without me having to try look for alternative ways of integrating it and making use of it in my own individual way
13.	I think other students/staff/users would find this application usable and useful
14.	Having the application as part of teaching this course, or learning this course, frustrates me
15.	It is somewhat frustrating to access information and download resources using the application
16.	Based on your overall experience with the use of other software and websites, how would you rate this application?
17.	I thought about referring to help and documentation during my use of the application
18.	I experienced errors during my use of the application
19.	If you experienced errors during your use of the application, what specific kind of errors did you experience?
20.	I felt somewhat apprehensive about using the application
21.	I am not sure how this application, its features and its content are meant to integrate with my learning or teaching
22.	The application makes me complete irrelevant , superfluous or distracting tasks that are not related to my learning (or teaching) processes

23. I do not believe that based on my opinions any serious improvements will be applied to the application
24. I can conclude that the institution does not understand my unique learning or teaching approach and learning or teaching processes
25. I should not be asked to modify my learning and/or teaching approach in order to work effectively with this technology application
26. My use of the application was guided by training aids and technical documentation
27. Some people say that eLearning applications are just tools that present existing information, and are of little value in student construction of new knowledge, collaboration and deep learning. Based on your experiences with this application, would you agree?
28. Some people say that eLearning applications are valuable tools that encourage the active collaboration between students and staff, exploration of different opinions and promote a culture of social learning. Based on your experiences with this application, would you agree?
29. Some people say that eLearning applications enable the effective acquisition of practical knowledge and skills that students can then apply in real life situations. Based on your experiences with this application, would you agree?
30. I believe I understand, and am well informed of the overall way this application, its features and its content should integrate with student learning and teaching
31. Even if an application does not integrate well into my learning or teaching approach and learning or teaching processes, it is important that it is flexible enough so that I can integrate and use it in my own individual way
32. This application is not flexible enough to integrate with my learning and teaching approach

The five sub-scales that are applied to the sub-layer 1 instrument, which assist to establish linkages to sub-layers 2 and 3 are shown in Table 5:

Table 5 – Sub-factors Identified for Establishing Links to Sub-Layers 2 and 3

Factor	**Related Survey Items**
Usability	13, 15, 16, 19, 10, 25, 28, 11, 27
Feelings toward an application (emotional aspects of technology evaluation)	18, 26, 21, 12, 14
Resistance to implied authorities in relation to one's teaching and learning	29, 30, 23, 8
Preconceived notions about eLearning	6, 7
Flexibility of application in relation to the teaching and learning	24, 22

Survey instruments for sub-layers 2 and 3 have also been developed, but there is insufficient room in this chapter to accommodate an adequate description of the work that was put into designing these instruments. A full overview of all sub-layer instruments can be found in Quinton and Logunov (2010). Table 6 provides a brief outline of the expected strategic outcomes to be derived from sub-layers 1, 2, and 3. Tables, 7, 8, and 9 provide additional detail on the key components and evaluation methods that are applied to sub-layers, 1, 2, and 3.

Table 6 – **Synopsis of Tables 7, 8, and 9 showing Strategic Outcomes**

Sub-layer	**Parameters for Surveys and Analysis**	**Strategic Outcomes**
1	Generic usability, usefulness, and educational value (UUE) –> technology effectiveness for educational use	Recommendation of technology for inclusion/non-inclusion into the TELT Platform based on UUE parameters -> Strategies for evaluating and determining educational value of technologies.
2	Application Specific UUE Evaluation and Application Specific Student Attitudes,	Strategies for integrated learning behaviour and effective TELT application

	Expectations and Perceptions Assessment -> learning effectiveness of the TELT application based on: • integration of knowledge in the learner • integrated development of learners	implementation that lead to enhanced learning effectiveness in online L&T.
3	Usability, Usefulness and Educational Value as applied to L&T environment-specific evaluation -> learning effectiveness of an Integrated L&T/TELT environment based on Expectations (E), Perceptions (P), Reactions (R) , and Conclusions (C) – EPRC analysis.	Strategies for Integrated Learning Behaviour and Effective Educational Technology Strategy Strategies for effective institutional diffusion, integration, and acceptance of TELT innovations.

Note, for a detailed description and explanation of the terms ‘ educational value’, ‘usability’, ‘usefulness’, and ‘learning effectiveness’, refer to Quinton and Logunov (2010). Note also, that the pages numbers provided in Tables 6, 7, and 8 once again refer to the document Quinton and Logunov (2010). In brief however, an educational technology selected for Stage Three pilot testing on the TELT platform is evaluated on the basis of its educational value, usefulness, and usability for learning and teaching. For the purposes of this chapter, these three important terms are defined as:

- Educational value: a term to indicate the integration of knowledge in the learner [sic] – that is, the transmission of knowledge to the learner) and the integrated development of knowledge by the learner (Crittenden, 1968).
- Usefulness: a term that applies to the capacity to apply pedagogical strategies to effect a positive or improved learning outcome (that is, learning effectiveness).
- Usability: a term to describe the ease to which users can perform tasks on an educational application or solution to achieve particular learning goals associated with those tasks.

Although the terms usefulness and usability appear similar, there are several distinctions that have been observed in the construction of the evaluation framework:

- Usefulness informs educational rationales, while usability informs the usage of the software per se, that is, the usage may or may not reflect educational purposes.
- The measurement of usefulness requires the lecturer’s and students’ participation whereas usability may be evaluated by experts who may not be educators.
- If the technology application is not designed according to educational principles, its educational usefulness is low, but the usability could still be high as long as the users feel it easy to perform the tasks they are expected to perform.

STAGE 3 – TELT EVALUATION FRAMEWORK GOAL 1 (TABLE 7)

Evaluating Educational Value and Usefulness of Technology for Inclusion in the TELT Platform

USABILITY + USEFULNESS + EDUCATIONAL VALUE ➔ TECHNOLOGY EFFECTIVENESS

SUB-LAYER 1

GENERAL, QUANTITATIVELY FOCUSED EVALUATION OF TELT TECHNOLOGIES

DEMOGRAPHICS

- **Respondents are classified as Students and Staff**
- **Respondents are grouped by their Attitudes Towards Technology** (anti-bias measure)
- **Respondents are grouped by the self-assessment of the degree of opinions** (anti-bias measure)
- **Respondents are invited to contribute short Qualitative feedback on any of the Evaluation Questions**

① **Usability, Usefulness and Educational Value (UUE) of eLearning Technology** (Generic Evaluation) *page 18*
Usability, Usefulness and Educational Value are evaluated using instruments based on an extensive literature review of eLearning usability attributes, usability experience factors and educational value determinants

Quantitative Tool

Generic UUE Evaluation: ***page 24***
A sample of questions should representatively cover the six stages of the UUE life-cycle to include

CURRENT UUE STAGES:

- User Expectations
- User Experience
- On-Going Interaction
- Learning and Teaching
- Support
- Subjective User Satisfaction

➔ **UUE**

Student Attitudes, Expectations and Perceptions Assessment:
Additional questions are included to reduce bias for objective evaluation feedback, exploring dynamic factors for a better understanding of students and staff using the TELT application: understanding their learning and teaching needs, processes and attitudes.

CURRENT DYNAMIC FACTORS:

- **Feelings toward an application (emotional aspects of technology evaluation)**
  (anti-bias measure separates out emotion from Generic UUE Evaluation to provide greater objectivity)
- **Resistance to implied authorities in relation to one's teaching and learning**
  (anti-bias measure separates out personality traits from Generic UUE Evaluation to provide greater objectivity)
- **Preconceived notions about eLearning**
  (anti-bias measure separates out emotion from Generic UUE Evaluation to provide greater objectivity)
- **Flexibility of application in relation to the teaching and learning**
  (additional questions following on from UUE (Learning and Teaching) and UUE (Subjective User Satisfaction)

▼

RECOMMENDATION OF TECHNOLOGY FOR INCLUSION/NON-INCLUSION INTO THE TELT PLATFORM BASED ON EDUCATIONAL USABILITY, USEFULNESS AND EDUCATIONAL VALUE

▼

STRATEGIES FOR EVALUATING AND DETERMINING EDUCATIONAL VALUE OF TECHNOLOGIES

STAGE 3 – TELT EVALUATION FRAMEWORK GOAL 2 (TABLE 8)

Integrating TELT Applications for Learning Effectiveness

USABILITY + USEFULNESS + EDUCATIONAL VALUE → LEARNING EFFECTIVENESS OF THE TECHNOLOGY

SUB-LAYER 2

DETAILED, QUANTITATIVE AND QUALITATIVE EVALUATION OF LEARNING EFFECTIVENESS OF TELT APPLICATIONS

DEMOGRAPHICS	• Respondents are classified as Students and Staff • Respondents are invited to contribute short Qualitative feedback on any of the Questions

① **Usability, Usefulness and Educational Value of eLearning Technology** (Application-Specific Evaluation) ***page 18***

Quantitative Tool	• **Application Specific UUE Evaluation** • **Application Specific Student Attitudes, Expectations and Perceptions Assessment**

② **Learning Effectiveness Components** (In-depth Pedagogical Dimensions) ***page 62***

Qualitative Tool	**Selected Technology should support the following processes, course/educational design:**
COMPONENT 1 **Integration of knowledge in the learner** *Crittenden (1968)*	• **Does the learning system allow for cognitive level and competence level based progressive learning?** • **Assessment of the relationships between respondent's learning, their course instructor and the quality of course content (***Ozkan and Koseler, 2009)* • **Assessment of a respondent's perceived learning**
COMPONENT 2 **Integrated development of learners** *Crittenden (1968)*	• **Does the learning system facilitate community building / social learning?** o **Facilitation of interaction with the Teacher/Instructor/Support Staff** o **Facilitation of interaction with student peers** • **Does the learning system facilitate knowledge building**

↓

LEARNING OUTCOMES

Quantitative and Qualitative Tools	**Quantitative Online/Offline Surveys (similar to Sub-Layer 1, but more comprehensive)**

↓

STRATEGIES FOR INTEGRATED LEARNING BEHAVIOUR AND EFFECTIVE TELT APPLICATION IMPLEMENTATION INTO L&T

→ENHANCED LEARNING EFFECTIVENESS

STAGE 3 – TELT EVALUATION FRAMEWORK GOAL 3 (TABLE 9)

Evaluating Learning Effectiveness of TELT Application Integration into L&T Practice

USABILITY + USEFULNESS +	EDUCATIONAL VALUE ➔	LEARNING EFFECTIVENESS OF THE LEARNING ENVIRONMENT

SUB-LAYER 3

SPECIFIC, HIGHLY DETAILED, QUALITATIVELY FOCUSED EVALUATION OF INTEGRATED L&T / TELT ENVIRONMENTS

DEMOGRAPHICS

- Respondents are classified as Students and Staff
- Respondents are invited to contribute detailed Qualitative feedback on any of the Questions

① **Usability, Usefulness and Educational Value** (L&T Environment-Specific Evaluation) *page 18*

Quantitative & Qualitative Tools	**Evaluation of a Learning Environment Supported by TELT** In-depth, context-specific questions

② **Learning Effectiveness of an Integrated L&T/TELT Environment** *page 60*

Qualitative Tools

Determine integrated learning effectiveness by using an in-depth assessment to explore staff and student learning and teaching processes, and the impact and integration of L&T with the learning environment, course content, Learning and Teaching Processes (LTP), and Educational Value

Expectations (E) →	Perceptions (P) →	Reactions (R) →	Conclusions (C) →
Explore Student Understanding of learning Create Relevant Context for Student Responses Explore and Eliminate Potential Biases Understand Ways in which Technology is Used		**Integration of Knowledge, and Learner Development** Explore and Evaluate the Actual Changes in Integrated Student LTP	**Evaluate Educational Value**

THE "EPRC" PROCESS

- **Iterative Process:** the instrument informs itself of changing student views and associated biases, and is thus able to iteratively examine learning processes with increased scrutiny, relevance and detail
- **Contextual Process:** by coming to a mutual acknowledgement of the aims of the evaluation and an understanding of L&T functionalities and student integration, the process readily lends itself to contextual variations and changes in research scope
- **Qualitative Analysis:** facilitates in-depth exploration of real-life student and staff experiences with TELT Integration and individual Learning and Teaching Processes, backed up by information collected on their attitudes and approaches to technology and education through the EPRC process

Strategies for Integrated Learning Behaviour and Effective Educational Technology Strategy *page 64*

Quantitative and Qualitative Tools	**Online Learning Participation** **Learning Conceptions** **Student Profiling Strategy** **Usage Patterns and Engagement** **TELT Learning Environment Evaluation**	**Learning Environment design strategies and learning effectiveness (*pages 72-74*)** Informed by, and informs: INTEGRATED CURRICULUM DEVELOPMENT / EDUCATIONAL DESIGN / TECHNOLOGY IMPLEMENTATION APPROACH

▼

LEARNING OUTCOMES

Quantitative and Qualitative Tools	**Qualitative Individual and Small-Group Focus Groups and Studies** **Student Grades, Results, and Staff and Student Attitude Trends; Multi-Period Analyses**

▼

EVALUATION OF INTEGRATED LEARNING ENVIRONMENT / TELT EFFECTIVENESS -> STRATEGIES FOR EFFECTIVE INSTITUTIONAL DIFFUSION, INTEGRATION, AND ACCEPTANCE OF TELT INNOVATIONS (page 75)

CONCLUSION

An extensive literature review conducted over several months confirmed the difficulty in identifying a comprehensive, generic approach and methodologies suitable for facilitating the selection and evaluation of the effectiveness of educational technology applications and solutions that focus exclusively on serving the needs of learning and teaching. In the UNSW context, a well designed, highly flexible evaluation framework is crucial to ensuring all software applications and tools selected for the TELT platform will meet the demands of learning and teaching both now and in the future.

In summarising the outcomes of the review, consideration has been given to the variety of technological tools employed by many educational institutions and researcher and the wide use of those tools as determined by each institutional context. The present report provides a critique of available educational technology evaluation models and identifies the importance of certain factors that have been omitted in the evaluative models that were the subject of the review.

Based on the results of the review, a number of factors required further examination to assist in formulating a generic evaluation framework that affords an in-depth analysis of how to apply established learning and teaching principles to divergent contexts while integrating educational technologies to enhance the practical day-to-day application of these principles. The key factors identified for devising a viable framework include:

- synthesise the major literature findings regarding the role of the pedagogical, technological, and business variables as applied to the evaluation process
- avoid a narrow evaluative scope, but facilitate "scaling down" to the especial needs and preferences of particular departments or schools
- develop a fully inclusive, multiple stakeholders and multiple-values evaluation model where evaluation directly aligns to the timing of an evaluative activity
- apply an iterative, formative evaluation cycle (a feedback-improvement cycle)
- apply a descriptive, participatory evaluation approach
- employ multiple methodologies to assess each technological element and integrate learners' participation into the evaluative process, and
- broaden the scope of evaluation to include a core set of educational technology tools and "worldware".

The systematic / systemic structure of the TELT evaluation framework and the supporting evaluation methodologies serve to underpin and are underpinned by the UNSW vision and guidelines, and the L&T Enhancement Plan 2011 – 2015. In developing the framework, the overarching goal was to devise a 'living' evaluation framework that affords an iterative refinement process to underpin: all TELT technology selections and developments; efficient tracking of the trajectory of success (and failure) of educational technologies used across faculties and over time; and, ultimately the enhancement of the divergent learning and teaching approaches and practices in online learning at UNSW.

Finally, consistent with the notion of a 'living evaluation framework', several subsequent refinements have been conducted over the past two years, which have resulted in the establishment of a robust and highly informative approach to the selection and the assessment of the educational value and learning effectiveness of educational technologies currently in use at UNSW. For a complete descriptions and explanations of the ongoing application and refinement of the TELT Evaluation Framework go to:

http://telt.unsw.wikispaces.net/Evaluating+TELT

REFERENCES:

Agostinho, S, Oliver, R., Harper, B., Hedberg, H., and Wills, S. (2002). *A tool to evaluate the potential for an ICT-based learning design to foster "high-quality learning".* In A. Williamson, C. Gunn, A. Young., & T. Clear (Eds.), Winds of change in the sea of learning. Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, 29-38, Auckland, New Zealand: UNITEC Institute of Technology.

Ainley, J., Banks, D., and Fleming, M. (2002). The influences of IT: Perspectives from five Australian schools. *Journal of Computer Assisted Learning*, 18, 395–404.

Alden, J. (1983). *Manager training systems.* Report from senior HRD mangers forum on "Measuring HRD". Stamford, CT: Xerox Corporation.

Alexander, S. (1999). An Evaluation of Innovative Projects Involving Communication and Information Technology in Higher Education. *Higher Education Research & Development*, 18(2), 173-183.

Alexander, S. and Golja, T. (2007). Using Students' Experiences to Derive Quality in an eLearning System: An Institution's Perspective. *Educational Technology & Society*, 10 (2), 17-33.

Alexander, S., and Hedberg, J. (1994). Evaluating technology-based learning: Which model? In K. Beattie, C. McNaught & S. Wills (Eds), *Multimedia in higher education: Designing for change in teaching and learning.* 233-244, Amsterdam: Elsevier.

Anderson, D. and Johnson, R. (2006). *Ideas of leadership underpinning proposals to the Carrick institute: a review of proposals from the 'leadership for excellence in teaching and learning program'.* Carrick Institute. Retrieved 28 May, 2010, from http://www.altc.edu.au/resource-ideas-of-leadership-underpinning-proposals-altc-2006

Arnold, S. (2003). Content management's new realities. *Onlinemag*, 27 (1). Retrieved from http://www.infotoday.com/Online/jan03/arnold.shtml

Attwell, G. (Ed.) (2006). *Evaluating eLearning: a guide to the evaluation of eLearning.* Evaluate Europe Handbook Series, 2, Bremen, Germany.

Bach, J. (1994). The immaturity of CMM. *American Programmer*, *7*(9), 13-18.

Baker, E. and Herman, J. (2003). A distributed evaluation model. In G. Haertel & B. Means, *Evaluating educational technology: effective research designs for improving learning*, 95-120, New York, NY: Teachers College Press.

Batpurev, B. (2005). *Evaluation and adaptation of Open Source software for distance learning in Asia*. Paper presented at the 19th AAOU annual conference, Jakarta September 15-17, 2005. Retrieved on November 15, 2009 from http://www.pandora-asia.org/downloads/05-AAOU_Batpurev.pdf

Baumgartner, P. (1995). Didaktische Anforderungen an (multimediale) Lernsoftware. In: L.J. Issing, P. Klimsa (eds.) *Information und Lernen mit Multimedia*, Weinheim: Beltz, 241-252

Baylor, L. A., and Ritchie, D. (2002). What factors facilitate teacher skill, teacher moral, and perceived student learning in technology-using classroom? *Computers and Education*, 39, 395–414.

Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S. and Reeves, T.C. (2002). *Usability and Instructional Design Heuristics for E-Learning Evaluation*. In P. Barker & S. Rebelsky (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002. 1615-1621, Chesapeake, VA: AACE.

Benvic Project (2002). *Evaluation Methodology report: Benchmarking of virtual campuses*. Retrieved on October 30, 2009 from http://www.benvic.odl.org/16_02_tot.pdf

Benvic Project (2000). *Benchmarking of virtual campuses*. Retrieved on November 7, 2009 from http://www.benvic.odl.org/deliverable2.pdf

Boud, D., and Prosser, M. (2002). Appraising new technologies for learning: a framework for development. *Educational Media International*. 39(3), 237- 245.

Bowman, D., Gabbard, J., and Hix, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence*, 11(4), 404-424.

Boys, J. (2002). *Managed Learning Environments, joined up systems and the problems of organizational change.* Report to JISC Building MLEs in Higher Education Programme.

Britain, S., and Liber, O. (1999). *A framework for pedagogical evaluation of virtual learning environments.* Report to JISC Technology Applications Programme. Retrieved October 3, 2009 from http://www.leeds.ac.uk/educol/documents/00001237.htm

Britain, S. and Liber, O. (2004). *A framework for the pedagogical evaluation of elearning environments.* Report to JISC. Retrieved October 31, 2009 from http://www.cetis.ac.uk/members/pedagogy/files/4thMeet_framework/VLEfullReport

Bruce, B. (1999). Challenges for the evaluation of new information and communication technologies. *Journal of Adolescent & Adult Literacy, 42*. Retrieved September 20, 2009 from http://www.readingonline.org/electronic/jaal/March_Column.html

Bruce, B. C., and Rubin, A. D. (1993). *Electronic Quills: A situated evaluation of using computers for writing in classrooms.* Hillsdale, NJ: Lawrence Erlbaum.

Carmean, C., and Brown, G. (2005). Measure for measure: assessing course management systems. In P. McGee, C., Carmean, and A. Jafari, (Eds.), *Course management systems for learning: Beyond accidental pedagogy,*1-14, Hershey, PA: IDEA Group Publishing.

*Chalmers, P. A. (2003). The role of cognitive theory in human-computer interface. *Computers in Human Behavior,* 19(5), 593-607.

Chen, H.-T. (1990). *Theory-driven evaluations.* Newbury Park CA: Sage publications.

Cheon, J., and Grant, M.M. (2009). Are Pretty Interfaces Worth the Time? The Effects of User Interface Types on Web-Based Instruction. *Journal of Interactive Learning Research*, 20(1), 5-33.

Clark, R. E. (1994). Assessment of distance learning technology. In E. Baker & H. O'Neil, *Technology assessment in education and training*, 63-78, Hillsdale, NJ: Lawrence Erlbaum.

Chickering, A.W., and Ehrmann, S.C. (1996). Implementing the seven principles: Technology as lever. *AAHE Bulletin, 49*, 3-6.

Conole, C. (2004). E-Learning: the hype and the reality. *Journal of Interactive Media in Education*. Retrieved October 7, 2009 from http://www-jime.open.ac.uk/2004/12/conole-2004-12.pdf.

Cousins, J.B., and Whitmore E. (1998). Framing participatory evaluation. In E. Whitemore (Ed.), *New directions for evaluation,* 80, 5-23, San Francisco, CA: Jossey-Bass.

Cox, E.P. (1980). The optimal number of response alternatives for a scale: A review. *Journal of Marketing Research*, 17(4), 407–422.

Crittenden, B. (1968). What is Educational Value? *The Journal of Value Inquiry*, 2(4), 235.

Crowther, M.S., Keller, C.C., and Waddoups, G.L. (2004). Improving the quality and effectiveness of computer-mediated instruction through usability evaluations. *British Journal of Educational Technology*, 35, 289–303.

Crozat, S., Hû, O., and Trigano, P. (1999). *A method for evaluating multimedia learning software*. In Proceedings of the IEEE international Conference on Multimedia Computing and Systems - Volume 2 (p. 9714). ICMCS. IEEE Computer Society, Washington, DC.

Cunningham, S., Ryan, Y., Stedman, L., Tapsall, S., Bagdon, K., Flew, T., Coaldrake, P. (1998). *New media and borderless education: review of the convergence between global media networks and higher education provision.* Department of Employment, Education, Training and Youth Affairs, Evaluations and Investigations Program, Higher Education Division. Canberra: Australian Government Publishing Service.

Curriculum Corporation (2004). *Framework for evaluating online learning systems.* Retrieved on September 21, 2009 from http://api.edna.edu.au/attachment/2150/21577/1/elearn_environment_framework.doc

Dean, R., Biner, P., and Coenen, M. (1996) Distance education effectiveness: A system approach to assessing the effectiveness of distance education. *Education at a Distance: ED, 10*(3), J1-J4.

Dublin, L. (2003). If You Only Look Under the Street Lamps…or Nine e-Learning myths. *The e-Learning Developer's Journal*, June, 1-7.

Ehrmann, S. (1995). Asking the right question: what does research tell us about technology and higher learning? *The Magazine of Higher Learning*, XXVII (2), 20-27.

Franklin, N., Yoakam, M., and Warren, R. (1996). *Distance learning: a guide to system planning & implementation.* Bloomington, IN: School of Continuing Studies, Indiana University.

Freire, P. (2007) *Pedagogy of the Oppressed*. New York: Continuum International Publishing Group.

Graf, S. and List, B. (2005). *An Evaluation of Open Source ELearning Platforms Stressing Adaptation Issues*. In Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies, 163-165, Taiwan: IEEE Press.

Graham, C., Cagiltay, K., Lim, B.-R., Craner, J., and Duffy, T. (2001). Seven principles of effective teaching: a practical lens for evaluating online courses. *Technology Source*. Retrieved October 19, 2009 from http://sln.suny.edu/sln/public/original.nsf/0/b495223246cabd6b85256a090058ab9

Gunn, C. (1999). They love it, but do they learn from it? Evaluating the educational impact of innovations. *Higher Education Research & Development*, *18*(2), 185 – 199.

*Haag, B. B., and Snetsigner, W. (1993). *Aesthetics and screen design: An integration of principles*. Proceedings of the 25th Annual Conference of the International Visual Literacy Association,

92-97. Retrieved June 30, 2006 from http://eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/26/9e/f4.pdf

Hagner, P.R. (2001). *Interesting practices and best systems in faculty engagement and support*. Paper presented at the National Learning Infrastructure Initiative Focus Session, Seattle, USA.

*Hannafin, M. J., and Hooper, S. (1989). An integrated framework for CBI screen design and layout. *Computers in Human Behavior*, 5(3), 155-165.

Hedberg, J., Wills, S., Oliver, R., Harper, B. and Agostinho, S. (2002). *Developing Evaluation Frameworks for Assessing Quality ICT-based Learning in Higher Education*. In P. Barker & S. Rebelsky (Eds.), Proceedings of ED-MEDIA 2002, World Conference on Educational Multimedia,

Hom, James. (1998). *The Usability Methods Toolbox Handbook*. Updated 22nd August, 2003. Retrieved from http://jthom.best.vwh.net/usability/usable.htm

*Hron, A. (1998). Metaphors as didactic means for multimedia learning environments. *Innovations in Education and Training International*, 35(1), 21-28.

Institute for Higher Education Policy (2000). *Quality on the line: benchmarks for success in internet-based distance education. Study commissioned by the national educational association and blackboard inc.* Retrieved on October 15, 2009 from http://www.ihep.org/assets/files/publications/m-r/QualityOnTheLine.pdf

*Khentout, C., Harous, S., Douidi, L., and Djoudi, M. (2006).Learning and navigation assistance in a hypermedia. *International Journal of Instructional Media*, 33(3), 265-276.

Kiili, K. (2005). Content creation challenges and flow experience in educational games: The IT-Emperor case. *The Internet and Higher Education*, 8(3), 183-198.

Kim, S. Douglas, M. B. Orkand, A. and Astion, M. L. (2001) Design implications from a usability study of GramStain-Tutor, *British Journal of Educational Technology*, 32(5), 595-605.

Kirkpatrick, D.L. (1983). Four steps to measuring training effectiveness. *Personnel Administrator,* November 1983, 19-25.

Kozma, R. (1994) A reply: Media and method. *Educational Technology Research and Development, 42(2)*, 7-19.

Lam, P. and McNaught, C. (2004). *Evaluating educational websites: a system for multiple websites at multiple universities*. In L. Cantoni & C. McLoughlin (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications *2004*, (pp. 1066-1073). Chesapeake, VA: AACE.

Learning Technology Dissemination Initiative (1998). *Evaluation cookbook.* Retrieved on October 10, 2009 from http://www.icbl.hw.ac.uk/ltdi/cookbook/contents.html

Lefoe, G., Gunn, C. and Hedberg, J. (2002). Recommendations for teaching in a distributed learning environment: The students' perspective. *Australian Journal of Educational Technology*, *18*(1), 40-56. Retrieved on November 15, 2009 from http://www.ascilite.org.au/ajet/ajet18/lefoe.html

*Lyons, C. J. (2001).*Essential design for web professionals*. Upper Saddle River, NY: Prentice-Hall.

Marshall, S. (2007). *E-learning maturity model process descriptions.* Wellington, NZ: Victoria University of Wellington.

Marshall, S. and Mitchell, G. (2004). *Applying SPICE to e-learning: an e-learning maturity model*? In Proceedings of the sixth conference on Australasian computing education, 185-191, Dunedin, New Zealand.

McNaught, C. (2002). Adopting technology should mean adapting it to meet learning needs. *On the Horizon*, 10(4), 14-18.

McNaught, C., Phillips, P., Rossiter, D., Winn, J. (2000). *Developing a framework for a usable and useful inventory of computer-facilitated learning and support materials in Australian universities*. Evaluations and Investigations Program report 99/11. Retrieved October 11, 2009 from http://www.dest.gov.au/archive/highered/eippubs/eip99-11/developframework.pdf

Meiselwitz, G., and Sadera, W. (2006). *Usability and Student Learning Outcomes in Online Learning Environments*. In C. Crawford et al., (Ed.), Proceedings of Society for Information Technology & Teacher Education International Conference 2006, 448-453, Chesapeake, VA: AACE.

Mehlenbacher, B., Bennett, L., Bird, T., Ivey, M., Lucas, J., Morton, J., and Whitman, L. (2005). *Usable e-learning: A conceptual model for evaluation and design*. In Proceedings of HCI International 2005: 11th International Conference on Human-Computer Interaction, Volume 4 - Theories, Models, and Processes in HCI, 1-10, Las Vegas, NV: Mira Digital.

*Metros, S. E., andHedberg, J. G. (2002). More than just a pretty (inter) face: The role of the graphical user interface in engaging elearners. *Quarterly Review of Distance Education*, 3(2), 191-205.

Mills, D. (2005). *Future directions of course management systems*. In P. McGee, C. Carmean, A. Jafari (Eds.), Course management systems for learning: Beyond accidental pedagogy, 307-330, Hershey, PA: IDEA Group Publishing.

Moore, J.L., Dickson-Deane, C., Galyen, K., Vo, N. and Charoentham, M. (2008). *ELearning Usability Instruments What is being Evaluated?* In Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008, 430-435, Chesapeake, VA: AACE.

Nelson, C., Post, J., and Bickel, W. (2001). *Institutionalization of technology in schools checklist.* Retrieved October 12, 2009 from http://www.wmich.edu/evalctr/checklists/institutionalizationoftech.pdf

Nielsen, J. (1993). *Usability Engineering*. San Francisco: Morgan Kaufmann.

Nielsen, J. (1994a). Heuristic Evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods*, 25-62, New York: John Wiley & Sons.

Nielsen, J. (1994b). *Enhancing the explanatory power of usability heuristics*. In Adelson, B., Dumais, S., & Olson, J. (Eds.), Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence (CHI '94), 152-158.

Obexer, R. (2005). Deliverable 3: *LMS evaluation criteria & strategies.* Retrieved October 27, 2009, from http://www.tils.qut.edu.au/initiatives/learningmana/documents/LMS_Evaluation_Criteria_final_1.doc.

Oliver, M., McBean, J., Conole, G. and Harvey, J., (2002). Using a toolkit to support the evaluation of learning. *Journal of Computer Assisted Learning,* 18, 199-208.

Ozkan, S. and Koseler, R. (2009). Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation, *Computers & Education*, 53(4), 1285-1296.

Owen, J. (2006). *Program Evaluation: Forms and Approaches* (3rd Ed.). Sydney, Australia: Allen & Unwin.

Pangaro L. (2000). Investing in descriptive evaluation: a vision for the future of assessment. *Medical Teacher,* 22(5), 478-481.

*Parizotto-Ribeiro, R., and Hammond, N. (2005, July). *Does aesthetics affect the users' perceptions of VLEs?* Paper presented at the 12th International Conference on Artificial Intelligence in Education, Amsterdam, Denmark. Retrieved June 30, 2006 from http://www.informatics.sussex.ac.uk/users/gr20/aied05/finalVersion/RParizotto.pdf

Peng, L.K., Ramaiach, C.K, and Foo, S. (2004). Heuristic-Based User Interface Evaluation at Nanyang Technological University in Singapore. *Program: Electronic Library and Information Systems,* 38(1), 42–59.

Preston, C.C., and Colman, A.M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power and respondent preferences. *Acta Psychologica,* 104(1), 1–15.

Pohl, M. (2004). *Guidelines for E-Learning - Advantages and Disadvantages*. In L. Cantoni & C. McLoughlin (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004, 4827-4830, Chesapeake, VA: AACE.

Quinton, S, and Logunov, A. (2010). *A Multi-Definitional Approach to TELT Evaluation at UNSW: Stage 3 evaluation criteria – L&T.* Report prepared for the University of NSW Learning and Teaching Unit (L&T@UNSW). Retrieved from http://telt.unsw.wikispaces.net/Evaluating+TELT

Quinton, S. Pachman, M. and He, R. (2010). *Evaluation of the TELT platform: Essential elements and methodologies*. Report prepared for the University of NSW Learning and Teaching Unit (L&T@UNSW). Retrieved from http://telt.unsw.wikispaces.net/Evaluating+TELT

Raymond, E. S. (1998). The Cathedral and the Bazaar. First Monday, 3(3). Retrieved November 12, 2009 from: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/578/499/

Reeves, T. C., and Carter, B. J. (2001). *Usability testing and return-on-investment studies: Key evaluation strategies for web-based training*. In B. Khan (Ed.), Web-based training, 547-557, Englewood Cliffs, NJ: Educational Technology Publications.

Reeves, T., and Hedberg, J. (2008). *Evaluating ELearning: A User-Friendly Gu*ide. (in press).

Reeves, T., and Hedberg, J. (2003*). Interactive learning systems evaluation.* Englewood Cliffs, NJ: Educational Technology Publications.

Reeves, T., and Laffey, J. (1999). Design, assessment, and evaluation of a problem-based learning environment in undergraduate engineering. *Higher Education Research & Development, 18*(2), 219 – 232.

Roberts, T., Irani, T., Lundy, L., and Telg, R. (2004). Practices in student evaluation of distance education courses among land grant institutions. *Journal of Agricultural Education*, *45*(30), 1-10.

Ross, B. (2005). *Course management systems for learning: future designs*. In P. McGee, C. Carmean, A. Jafari (Eds.), Course management systems for learning: Beyond accidental pedagogy, 296-306, Hershey, PA: IDEA Group Publishing.

Saadé, R. G., He, X., and Kira, D. (2007). Exploring dimensions to online learning. *Computers in Human Behavior*, 23, 1721–1739.

Second Life (2010). http://secondlife.com/?v=1.1

Sim, G., Read, K., and Holifield, P. (2007). *Evaluating the User Experience in CAA Environments: What Affects User Satisfaction?* In Proceedings of the 10th International Computer Assisted Assessment Conference. Loughborough, UK.

Sim, G., Read, J.C. and Cockton, G. (2009). *Evidence based design of heuristics for Computer Assisted Assessment*. 12th IFIP TC13 Conference in Human-Computer Interaction, 204-216, Uppsala.

Siemens, G. (2004). Learning management systems: The wrong place to start learning. *Elearnspace*. Retrieved October 12, 2009 from http://www.elearnspace.org/Articles/lms.htm

Shneiderman, B., and Plaisant, C. (2004). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (4th ed.). Reading, MA: Addison-Wesley.

Stokes, P., and Stokes, W. (1996). Pedagogy, Power, Politics: Literate Practice Online. *Computers & Texts*, 13. Retrieved on November 19, 2009 from http://users.ox.ac.uk/~ctitext2/publish/comtxt/ct13/stokes.html

Stufflebeam, D.L., Foley, W.J., Gephart, W.J., Guba, E.G., Hammond, R.I., Merriman, H.O. & Provus, M.M. (1971) *Educational evaluation and decision making.* Itaska, IL: F.E. Peacock.

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

Sun Microsystems. (2001). *Sun usability labs and services: what is usability engineering?* Retrieved from: http://interface.free.fr/Archives/SUN_usability_benefits_Cost.pdf.

*Szabo, M., andKanuka, H. (1998). Effects of violating screen design principles of balance, unity and focus on recall learning, study time and completion rates. *Journal of Multimedia and Hypermedia*, 8, 23-42.

Technologies Enabled Learning and Teaching (TELT) (2009). *A Status Quo Report on the TELT Environment at July 2009.* Internal report. Learning and Teaching, the University of NSW: Sydney, Australia.

van den Berg, K. (2006). *Open source software evaluation*. In K. St. Amant & B. Still (Eds.), Handbook of research on open source software: technological, economic, and social perspectives, 197-210, Hershey, PA: IGI Global.

Venkatesh, V., Morris, M. G., Hall, M., Davis, G. B., Davis, F. D., & Walton, S. M. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Zammuto, R. (1982). *Assessing organizational effectiveness: systems change, adaptation, and strategy.* Albany, NY: SUNY Press.

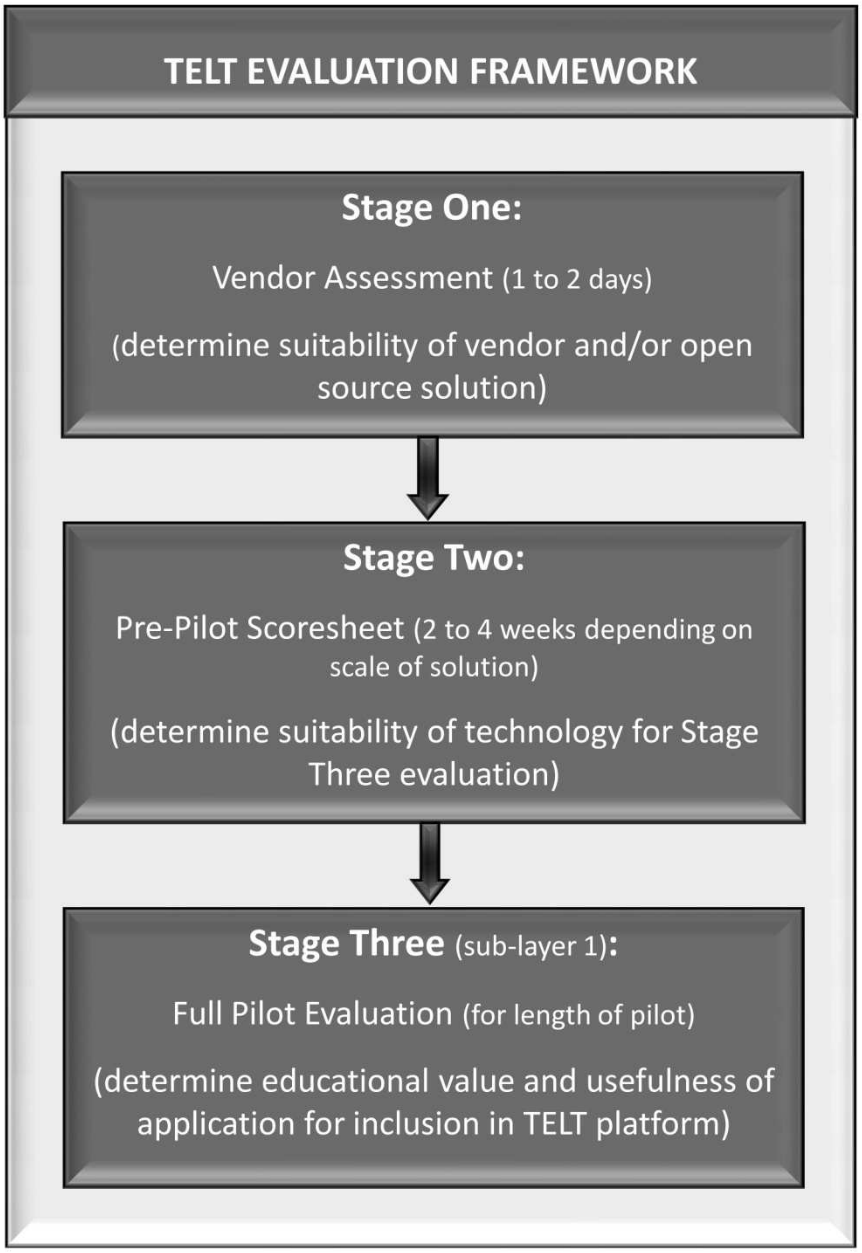


Figure 3: The three primary stages of the TELT Evaluation Framework

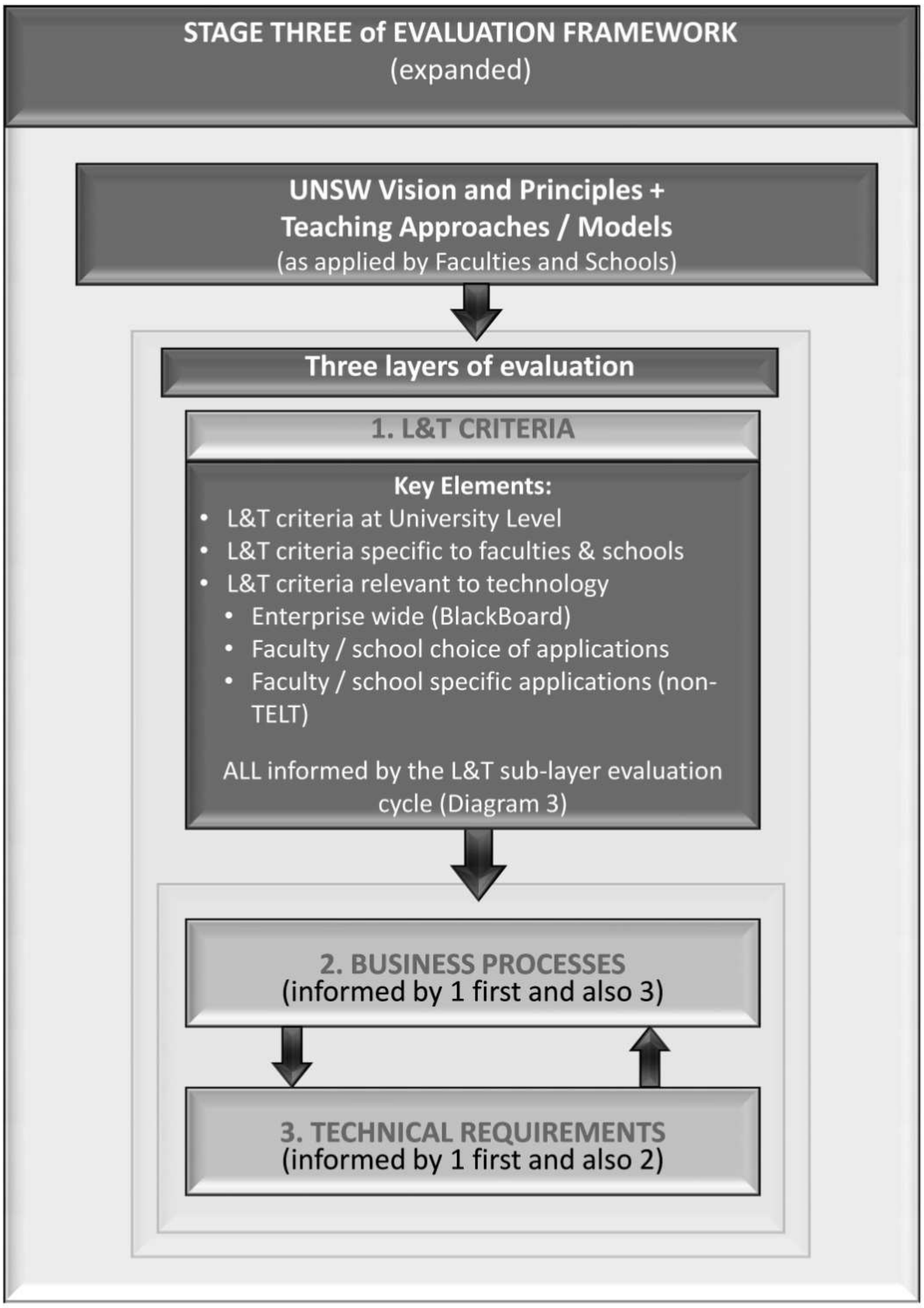


Figure 4: The three layers of evaluation in Stage 3

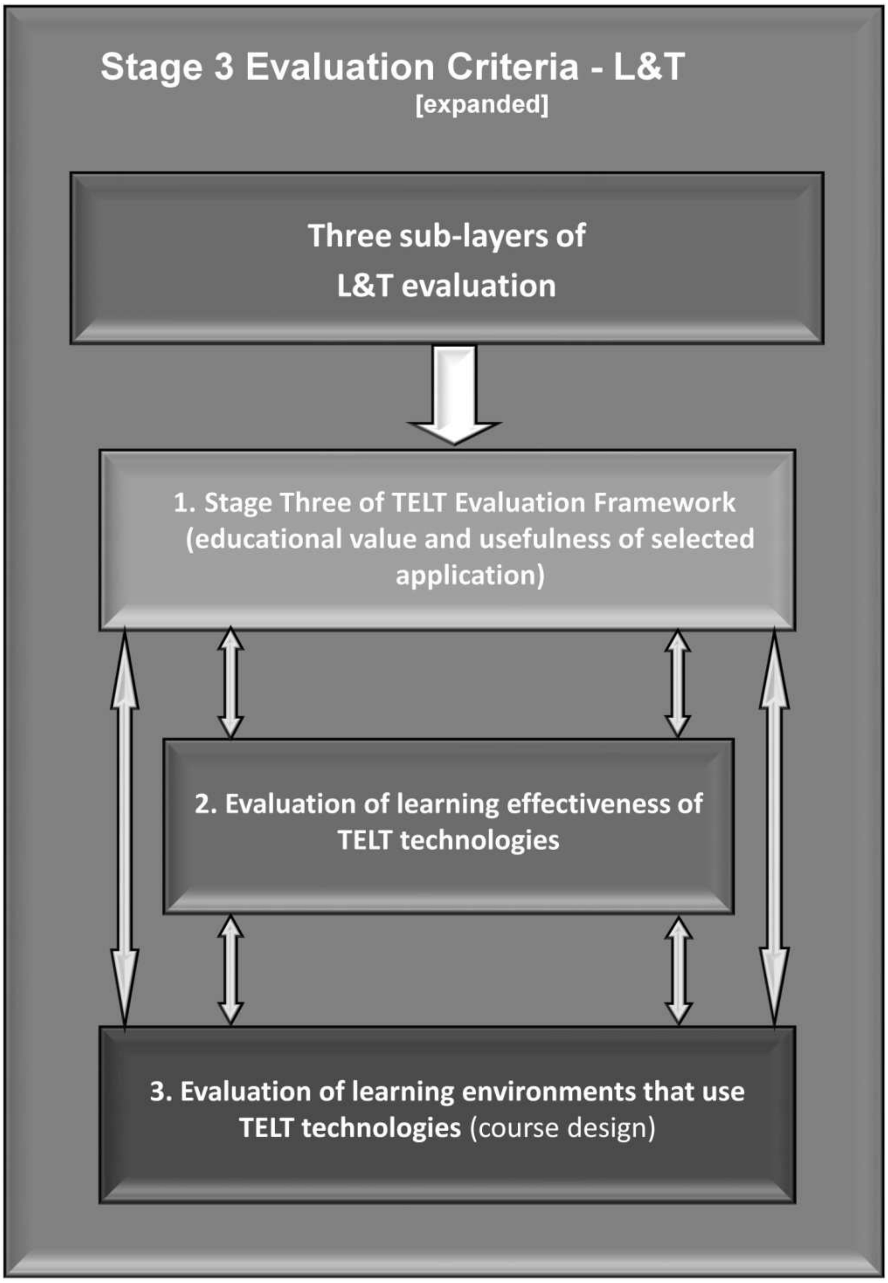


Figure 5: The three sub-layers of evaluation in Stage 3