Confidence in assessment decisions when using ICT

Kerry Earl
kearl@waikato.ac.nz
&
Bill Ussher
bussher@waikato.ac.nz
Faculty of Education, The University of Waikato, Hamilton, New Zealand.

Abstract

The central question addressed in this paper is: How can teachers and schools have confidence in their assessment decisions when using information communication technologies (ICT)? The answer centres on what makes quality assessment. Assessing and evaluating children’s achievement and progress is critical to development of sound curriculum programmes that focus on student outcomes. With the increasing use of ICT in schools and classrooms for a range of assessment purposes such as recording, data analysis and online activities, teachers and school leaders must be assessment capable in order to make informed decisions about assessment design, selection and modification that utilises ICT. Based on examining assessment purpose and the three principles of quality assessment (validity, reliability and manageability), this paper offers guidelines for classroom teachers, those with responsibility for student achievement and those who lead ICT policy and practice in schools to be critical consumers of ICT-based assessment tools, strategies and evidence. Vignettes of assessment practice using ICT are used to illustrate sound school and classroom practices in relation to validity, reliability, and manageability. Drawing from the work of assessment writers such as Crooks, Sutton, and Darr, the guidelines will assist teachers in the effective use of ICT for both formal and informal information gathering.
as well as for analysis and interpretation of information for summative and formative purposes. This knowledge is needed to underpin teacher confidence in their assessment decisions when using ICT towards ‘best fit’ for purpose.

Keywords: assessment-capable teachers; principles of quality assessment; school and classroom use of ICT.

Introduction

Increasingly information communication technologies (ICT) are being utilised to enhance teaching and learning in schools and classrooms (Lai & Pratt, 2007; Lim, 2007; Luterbach & Brown, 2011). In this article we use Earl and Forbes’ (2012) broad definition of ICT “as digital technology that allows us to search, assess, retrieve, store, organise, manipulate and present information” (p. 225). In spite of the growing potential of ICTs in transforming teaching and learning practices, Leeson and Hattie (2009) cautioned that this does not go far enough in elucidating complementary assessment practices, “some areas where technologies, whilst fully applied to learning approaches, have not yet been extended to the assessment of that learning” (p. 1). Literature reports a range of affordances of ICT, including students and teachers using ICT to create evidence of student learning, teachers and students using ICT to store evidence, ICT use in the reporting of evidence, and teachers using ICT to analyse evidence. There may be an attraction to using ICT for assessment purposes particularly regarding student engagement and aspects of manageability. However, Lim (2007) warns that, “When ICT is employed in schools, there is a need to review and modify existing teaching, learning and assessment practices; and hence, [an] education system must be responsive to adapt to these changes” (pp. 90-91). Studies indicate that where schools and teachers introduce ICT to enhance teaching, learning and assessment, there must be caution, creativity, collaboration and clear thinking in the planning, organising and use of ICT (Hammond, 2011; Lim, 2007; Underwood & Dillon, 2011). Teachers and school leaders need to have the knowledge to make informed decisions about assessment design, selection and modification when using ICT. The central question addressed in this paper
therefore is: How can teachers and schools have confidence in making assessment decisions when using ICT? The answer lies in what makes quality assessment.

**Evaluating assessment quality**

Assessment is a “complex process of gathering information about how learning is proceeding as it occurs” (Hill, 2012, p. 161). Effective assessment is “fit for purpose” meaning the outcome will provide sound information on which to make the desired decision (often regarding what happens next). Effective assessment is deliberate and articulate (clear, often recorded and questioned/reviewed). Assessment information is judged as sound according to the principles of reliability, validity and manageability (Darr, 2005 a&b; Sutton, 1992). These qualities also extend to the interpretation and use of assessment information. While acknowledging the importance of using a range of assessments, in this paper we focus on formal and informal assessment processes.

Teachers having the knowledge to make informed decisions about assessment activity design, selection and modification when using ICT will increase their confidence in making dependable assessment decisions. Assessment-capable teachers (Stiggins, 1991) therefore have the knowledge and ability to create and evaluate assessment tools and strategies to ensure the gathering of high quality evidence as the basis for well informed decisions about student progress, achievement, future learning and future teaching. This connection is important because we use analysed and interpreted assessment information to feed back to both teachers and students, to shape decisions and actions (such as resourcing decisions on learning assistance and special support), and to provide information for portfolios and reports, as evidence of mastery and for benchmarks, ranking and qualifications (see MoE, 2007, p. 40). Assessment-capable teachers use several criteria to guide decision-making about the quality of assessments. Crooks (1993) proposed the following four important assessment considerations and Hill restated these in 2012 (p. 175):

- “Will the assessment do any good?” [For example, does it have a clear and worthwhile purpose?]
• “Will the assessment do any harm?” [e.g. Will the experience impact negatively?]
• “Will decisions be based on a true and sufficiently broad picture?” [validity]
• “Will decisions be based on stable enough information?” [reliability]

In addition to the above four criteria, Sutton (1992) adds manageability as critical in the overall judgement of the value of assessment for learners and teachers. These five criteria are briefly explored. Keep in mind that in many ways one criterion may overlap with others and attempts to enhance the degree of one for best-fit in terms of purpose of an assessment often impacts on the degree of others and the balance. For example, by increasing the reliability of an assessment you may increase its manageability, but this is also likely to decrease the validity and perhaps impact on the extent that the assessment evidence will achieve its purpose.

**Will the assessment do any good?**

For assessment to do any good it must have a clear and worthwhile purpose. Possible purposes include meeting learning needs, celebrating achievement, determining progress, guiding decision-making (e.g. about support and use of resources) and motivating learners. Absolum, Flockton, Hattie, Hipkins and Reid (2009) emphasise the importance of assessments for affirming learning so that achievements can be celebrated, which in turn motivates learners. In order to determine progress, teachers need to have a clear idea of learning targets or outcomes represented in class curriculum and learning opportunities provided. Chappuis, Chappuis, and Stiggins (2009) point out, “if we don’t begin with clear statements of the intended learning – clear and understandable to everyone, including students – we can’t end up with sound assessments” (p.15).

Knowing who will use the results to guide decision-making about further learning will also help clarify purpose. Chappuis et al.’s (2009) proposal that effective communication of results connects to the worthwhile assessment purpose of meeting identified learning needs in a timely and understandable manner for those using results and providing
information on ‘next steps’ needs to be an important consideration. Chappuis and colleagues (2009) argue for student involvement in the assessment process. This serves to further emphasise that the design of assessment needs to motivate and support students in self-assessment and goal-setting in order that they can monitor their own progress and learning. This autonomy in assessing their own progress is necessary in sustaining confidence in their ability to learn over their lifetime. As Absolum et al. (2009) emphasised,

Students who have developed their assessment capabilities are able and motivated to access, interpret, and use information from quality assessments in ways that affirm or further their learning. Students cannot get there without help and support. To give the necessary support, many teachers may first need to strengthen their own assessment capabilities. (p. 19)

Assessment must do no harm

Teachers must inquire about the potential of assessment experiences to impact negatively on their learners. To minimise this risk, assessment-capable teachers ensure assessments, among other things, identify appropriate learning needs, provide appropriate opportunities to celebrate achievements publicly or privately, focus on desired knowledge and skills in order to accurately determine progress, gather sufficient evidence to confirm claims for additional learning support, and avoid any emotional responses that hinder learning or trigger disengagement (Absolum et al., 2009).

Validity: Decisions are based on a true and broad picture

Assessment-capable teachers determine that their decisions are based on a true and sufficiently broad picture, that is, the assessment information is valid. This is arguably the most important quality indicator for assessment. Darr (2005a) identifies that, “Validity should be at the top of our minds when we design assessments or make decisions about assessment programmes” (p. 56). The evaluation of validity is a matter of degree. No one assessment will be totally valid. Absolum et al. (2009) suggest the importance of validity to decision making, “is a function of both parts of the decision-
making process: if the descriptive part [scores, levels, observations, etc] is good but the prescriptive part [what to do next] is poor (in other words, if a student’s performance is correctly determined but the consequential decisions are detrimental to learning) or vice versa, the assessment lacks validity” (p. 34). Assessment-capable teachers are attentive to the ongoing utilisation of evidence gathered just as they are to the relationship between information gathered across a range of assessments when making decisions about further learning. For teachers to evaluate the degree of validity of a particular assessment requires professional judgment and consideration of all available evidence. This decision will be specific to particular context and relies on the teacher as decision-maker.

**Validity: Fairness**

Validity also concerns the equality or fairness of assessment design and content. A fair assessment allows students regardless of gender, ethnicity, religion, etc. to do equally well. All students should have equal opportunity to demonstrate the skills and knowledge being assessed. Some students should not have an advantage over other students completing the task. Questions teachers can ask to evaluate fairness are: is the assessment suitable for all students in the group or class? Will the results be unduly influenced by factors unrelated to the learning objectives or standards? Have all students being assessed had the same opportunity to learn?

**Reliability: Decisions are based on stable enough information**

Assessment-capable teachers determine that their decisions are based on stable enough information, that is, the evidence is reliable. Reliability is the extent to which an assessment is dependable across groups or administrations. The question assessment-capable teachers ask of an assessment to evaluate reliability is, “How dependable and consistent will the assessment results be across time, tasks and across markers?” (Darr, 2005b, p. 59). For example, the number and difficulty of tasks, the item types, and variations in marking can threaten reliability. Absolum et al. (2009) explained, “assessments are said to be reliable when inconsistency and error are reduced to a level
that is reasonable, considering the nature of the interpretations and decisions to be made” (p. 33). To decrease the chance of error and inconsistency and thereby increase reliability requires clear marking criteria, clear task instructions (for teachers and students), and the consideration of wellbeing of students to be assessed and test conditions to minimise chance factors. Harlen (2007. p.18) explained that reliability “refers to the extent that results can be said to be of acceptable consistency for a particular use” and goes on to emphasise that, “high reliability is necessary when the results are used by others and when students are being compared or selected” [emphasis in the original].

**Manageability**

Manageability is the extent to which assessment is user-friendly and efficient for learners and teachers (Sutton, 1992). In many ways this principle balances ideals with the practical considerations that determine what actually happens in assessment practice in classes and schools. Sutton (1992) explains that in “real situations in school, your aim of high-quality assessment procedures will inevitably be constrained by the resources at your disposal, of which the most crucial is your own time and energy. What you are aiming for is a ‘best fit’ model” (p. 17). Questions teachers can ask to evaluate manageability focus on the time, effort, resources and space required to develop and administer assessment effectively and efficiently. Whether the information to be gathered will be worth the effort is critical. Teachers need to consider preparation time as well as administration and marking time. The degree of manageability may also address the potential to provide prompt feedback, usability in a variety of situations, adaptability with minimal revision, usefulness of evidence for more than one purpose, and ease of analysis, manipulation and storage of the information.

**Vignettes of assessment using ICTs**

Increases in internet speeds and storage have expanded options for schools’ use of multimedia, computer adaptive technologies and virtual reality, including use for assessment. Leeson and Hattie (2009) suggested that, “where these technologies are being used their success is more measured in engagement than information gained”,

---

explaining that, “Only if the assessments are more dependable (valid and reliable) as well as more efficient (time, intrusiveness, immediacy, personalised) is it likely that these methods will replace many of the sturdy and well tried and known (usually pencil-and-paper) methods” (p. 12). What follows are three short examples of teachers using ICT in their assessment practice in ways that move beyond what is readily achieved using paper-and-pencil methods. The vignettes provide illustrative examples in answer to the question, ‘How can teachers and schools have confidence in their assessment decisions when using ICT?’ The teachers in the vignettes have made decisions about the quality of the design and use of assessment involving ICT that align with the purpose for which the information is to be gathered and used.

Performance assessment

This first vignette is an example of a teacher using technology to facilitate a performance assessment.

Shane often explores internet sites for material and activities to be used with his year 3 students (age 7). Recently his class was working on a unit on game playing. He set about teaching them to throw objects at a moving target in a health and physical education unit. Near the end of the unit Shane needed to assess the children’s achievements. Finding the activity titled Using Technology in Physical Education (http://vimeo.com/2361947) he decided that the clip was suitable to assess, for formative purposes, the children’s skill of throwing an object at a moving target. Shane projected the video clip on to the wall in the school hall, which provided the size and space required for the activity. He took advantage of the flexibility the technology allowed to make adjustments for some of the children. Using koosh-balls as the object, the children cycled through their turns at hitting the target from an appropriate distance. At times Shane observed the children throwing and provided feedback on improving their skill. At times the children worked in pairs and commented on each other’s skill. At other times Shane had assistance from a parent or student teacher to gather information and moderate observations and
judgments across markers as well as across sessions (assessment events). After 2–3 sessions Shane had enough evidence to make a reliable judgement about his students’ ability for summative purposes. Now, on some wet days, this activity is repeated for maintenance of this skill. Children frequently seek Shane’s observation of their progress.

This assessment activity using ICT included opportunities for both formative and summative assessment purposes. Using the same task for skill development and practice and for assessment purposes means a clear match resulting in high validity. The activity was adjusted to be appropriate for all the students in Shane’s class and so was a fair assessment. As Shane collected assessment information over time he increased the reliability of his judgments. Both Shane and his students had a clear understanding of what was being assessed. The consistency and thus reliability of this assessment was increased through the involvement of a range of people as markers and moderators. Because the school has the resources available in the hall, this was an accessible and manageable activity regardless of weather conditions. This activity is also easily reusable on a different occasion and with a different class with only a change in the recording sheet needed meaning the assessment meets the manageability criteria.

**Cognitive assessment**

Among the more popular assessment practices using ICT are those that utilise technology to elicit evidence of cognitive achievements and progress. While many internet sites offer online assessments, with claims of convenience and manageability in terms of saving you time to create, evaluate and analyse any assessment, the degree to which they are valid is a challenge. The complexity of matching the intended learning of a classroom programme with the content of an online assessment is often overlooked but it is critical if ‘next steps’ are to be determined. The next vignette highlights a teacher’s efforts to maximise validity and also the potential to assess for multiple purposes.
Chris believes in emphasising the development of expertise in her years 7 and 8 students (ages 11 to 13) through a significant themed study of one school term in length. Each year in assessing these studies for summative purposes, groups of students demonstrate an aspect of their learning by designing a 10-question quiz. Currently the quizzes are developed using the Quizstar site (http://quizstar.4teachers.org). Chris teaches the class how to develop a quiz for use with the next class that targets key learning goals. Chris can then use these quizzes as legacy items passed on from one class to the next. As the next class starts their study they complete a quiz to set up the expectations for their own learning and so this activity serves a formative function for the next class. Chris has been using this assessment for three years now over two different studies each year and the quizzes have a well-developed legacy. Students in following classes, particularly younger siblings, expect and look forward to this activity and the chance to demonstrate their expertise.

Both the process and the outcomes of the quiz development can be used for formative and/or summative purposes by both the class that develops the quiz and another class undertaking the same topic study. Chris’s activity asks children to understand the content and their learning in the selected study to the degree that they can act as teacher selecting what is worthwhile to assess, thus requiring both higher level thinking and a degree of metacognition. The use of software also can be challenging, as this requires a level of familiarity and skill in its use. If the children do not know the software well enough to be successful in completing the task then this assessment will not assess what it is designed to assess, significantly reducing validity. Chris is careful to provide instruction to ensure this is not the case. The reliability of the quizzes themselves is dependent on the item construction by the children, however because Chris uses them for formative purposes with the next class this is not a major issue. To ensure a high degree of reliability for summative purposes Chris assesses the quiz using a rubric negotiated with the children at the start of the summative phase.
This activity does take time. The children need to review their learning, learn the software and develop a quiz but once the children have familiarity with the software the activity can be reused with different content. A variety of assessment formats better enables our diverse students to show what they know. The development of the quiz by a whole class or even students in small groups puts demands on the hardware resources of the school. However, this activity is so motivating and supports such a higher level learning that Chris maintains it is worthwhile. Given the need for decisions about assessment quality to focus the validity and reliability of the evidence gathered for purpose Chris does need to ensure that each new ‘legacy’ quiz is suited to purpose, valid to the learning intentions of the current study and will cause no harm through embarrassing any individuals. However, as student groups change, units of work can readily be modified and adapted, so the form and content of quizzes can be updated to suit the new cohort and curriculum.

**Performance and cognitive assessment**

Continuing with the theme of children being ICT smart, this third vignette is an example of one teacher’s assessment literacy and ICT knowledge coming together to create an opportunity for assessment. Digital technologies can provide teachers and learners with opportunities to gather evidence of learning over time and space.

Children in Sam’s year 4 class began the year by producing portraits of themselves for the walls of their classroom. They then scripted and produced a three-minute video to introduce themselves and their interests using Aurasma. (Aurasma is an application that can embed a video in a static image that can be printed. It combines a GPS location and photo to allow additional media to be attached.) On their iPads the students link their videoed introductions to their portraits. Later in the year students used their iPads while at camp to collect video and stills of learning activities and evidence of their learning outcomes covering criteria that has been negotiated with their teachers. Each evening while they are at camp, a time is set aside for students to review their evidence and group reflective feedback is used to target next steps for enhancing their learning during
the following day. Once back in class, students turn this formative evidence into a multimedia summative product to share with parents by developing a sequence of still photographs and accompanying video, again using Aurasma (http://www.aurasma.com/).

The validity of the assessment from a strategy such as this depends on teacher and students having shared understanding of the specific learning goals when selecting the assessment evidence and recognising the information provided by the evidence, both day to day and in preparing a summative product. Using technology can tempt distraction from curriculum learning intentions and focus on impressing others with the use of technology and presentation. This was not the case here because the teacher and students had negotiated the success criteria. For an activity like this to succeed the issue of accessibility for all is paramount. In fact where this is not the case, social and emotional harm can occur. Others could have access to this assessment material for the purpose of moderation or reporting therefore the accessibility of this evidence can potentially enhance reliability. The manageability of the assessment evidence in this electronic application is also likely to be enhanced but will require software knowledge and skills in order to improve feedback and reporting of assessment outcomes.

These three vignettes represent the tip of the iceberg in terms of schools and teachers using ICT to enhance teaching, learning and assessment. However, they are examples that show the importance of the four considerations proposed by Crooks – *good, no harm, validity* and *reliability* for all students when gathering and using assessment evidence to promote learning.

**Conclusion**

In this paper we promote the idea that teachers can have confidence in their assessment decisions with ICT by evaluating the quality of an assessment using the three principles, validity, reliability and manageability, and remembering that for assessment to do any good it must have a clear and worthwhile purpose and do no harm to all students. The knowledge of the curriculum and individual learners that teachers bring to the assessment process will overlay their decisions about the quality of the assessment. We
have emphasised the importance of attention to the fitness for purpose of choices of
assessment and the use of assessment results for learning and decision-making. What
follows are some guidelines for classroom teachers, those with responsibility for student
achievement, and those who lead ICT policy and practice in schools, to use to support
them as critical consumers of ICT-based assessment, and to be confident in their
assessment decisions when using information communication technologies.

The following criteria are presented in six sections with questions to guide decision-
making regarding best use of ICT within an assessment programme. Many of the
questions apply to all assessment activities.

**Purpose**

1. Is this assessment truly fit for purpose?
2. Will this method of assessment yield the necessary information required for the
   purposes of those who are seeking to use it?
3. Will the assessment do no harm? That is, will the use of ICT or the assessment
   disadvantage, distress or demotivate any individual or group?

**Validity**

1. Is there is an unmistakable match with the intended learning from your unit of
   study?
2. Will this assessment provide sufficient information to be useful, relevant and
   valuable in your planning of further learning experiences?
3. Does the assessment assess knowledge and skills that are authentic to your children
   as learners and will apply to their life outside school?

**Validity (fairness)**
1. Are there any aspects of the assessment that need to be modified to avoid unnecessary bias associated with gender, language, ethnicity, culture or social background?

2. Is the content of the assessment meaningful to all those being assessed?

3. Does the assessment resource offer fair opportunity of successful completion by all students in terms of access to the ICT and ability to use the ICT?

**Reliability (marking)**

1. Are the marking criteria clear, easy to follow to enable consistency and dependability?

2. Are you able to check that the answers for the assessment are obvious and correct for your learners?

3. Does the marking schedule provide clear and unambiguous answers enabling consistent results and minimising potential for variations in marking?

**Reliability (items)**

1. Does this assessment provide clear, precise directions that are obvious, ensuring consistent interpretation by all students?

2. Is the assessment presented in a manner that is clear and concise, easy to follow and with language that is unambiguous and suited to the level of your learners?

3. Is the assessment of a suitable length for your learners to minimise chance factors and avoid concentration fatigue?

**Manageability (time and resources)**

1. Is this assessment an effective use of student, class and teacher time in terms of creation, modification, administration, marking and analysis for the evidence gathered?
2. Is the assessment easy to manage in a classroom situation, for example within time and required access constraints?

3. Is this assessment reusable – can it be easily modified to cater for another learning area, another time, or another cohort?

If the answer to any of these questions is, “After some adaptation or modifications”, then the question must follow: “Can you adapt or modify the assessment to meet this criterion to a sufficient degree?”

With the increasing use of ICT in schools and classrooms for a range of assessment purposes, teachers and school leaders must be assessment capable in order to make informed decisions about assessment design, selection and modification. There are many sites and solutions available on the internet to assist teachers with the assessment of student learning. Teachers must provide opportunities for learning with and through ICT within the classroom programme for any ICT to be an appropriate tool for assessment. Access for students to gain knowledge of software and the stability of infrastructure will influence the manageability of particular assessments using ICT. Inherent in this use of ICT are considerations of hardware and software issues. Students’ ability to demonstrate or report on what they know and have learned may also influence the appropriateness of the ICT in supporting assessment. In thinking about these issues teachers need to decide whether the ICT they have access to can purposefully contribute to the teaching, learning and assessment in their classroom and school programmes. This paper provides guidelines for teachers and for those who lead ICT policy and practice in schools to help them to be more critical consumers and advocates of ICT-based assessment and confident in their assessment decisions when using information communication technologies.

Acknowledgement

The authors wish to acknowledge the contribution of the members of the Curriculum and Assessment teaching team, Faculty of Education, University of Waikato, over a number of years in the development of this position and the guidelines presented.
References


This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivs 3.0 New Zealand License](http://creativecommons.org/licenses/by-nc-nd/3.0/nz/).