australasian journal of
TECHNOLOGY EDUCATION

The Australasian Journal of TechnologyEducation is a peer refereed journal, and provides a forum for scholarly discussion on topics relating to technology education. Submissions are welcomed relating to the primary, secondary and higher education sectors, initial teacher education and continuous professional development, and general research about Technology Education. Contributions to the on-going research debate are encouraged from any country. The expectation is that the Journal will publish articles at the leading edge of development of the subject area.
The Journal seeks to publish

- reports of research,
- articles based on action research by practitioners,
- literature reviews, and
- book reviews.

Publisher: The Technology, Environmental, Mathematics and Science (TEMS) Education Research Centre, which is part of the Faculty of Education, The University of Waikato, publishes the journal.

Contact details: The Editor, AJTE, pj.williams@waikato.ac.nz

Cover Design: Roger Joyce

ISSN: 2382-2007

australasian journal of
TECHNOLOGY EDUCATION

Criteria for Success: A study of primary technology teachers' assessment of digital portfolios

Eva Hartell
Inga-Britt Skogh

Abstract

Transparency regarding criteria for success in assessment processes is challenging for most teachers. The context of this study is primary school technology education. With the purpose to establish what criteria for success teachers put forward during the act of assessment, think-aloud protocols were collected from five primary teachers during an assessment act. Results are based on content analysis of think-aloud protocols and quantitative measures of reliability in order to ascertain teachers' motives for decision-making when assessing Year 5 pupils' multimodal e-portfolios.

Findings show consensus among these teachers, focusing on the execution of the task in relation to the whole, rather than to particular pieces of student work. The results confirm the importance of task design, where active learning in combination with active tutoring is an integral part, including provision of time and space for pupils to finish their work.

Keywords: assessment, technology education, primary school, e-portfolio, adaptive comparative judgments, teachers.

Introduction

The context of this study is primary school technology education and, in particular, technology teachers' assessment practices. Assessment conducted by teachers has long been, and still is, a high-profile issue for politicians, educational researchers and, not least, school leaders and teachers in Sweden. Assessment is a difficult art, it links teaching with learning and offers strong potential for pupils to further their learning. There are various forms of assessment and it can be conducted in different ways in order to fulfil different purposes (Harlen, 2012; Newton, 2007). Assessment includes at least three basic parameters. One is the definition and delineation of what is to be assessed (the required qualities 'in themselves' but also the shape/way that these qualities are portrayed). Another basic parameter is the assessment's subject location, that is, the individual/s who make the assessment and the knowledge, experience, and judgment (s)he/they possess. A third fundamental parameter is the object of assessment, which is, how and to what extent the qualities to be judged are expressed and embodied by/through the object of assessment (e.g. pupils' presentations, statements, and models). Assessments could and should embrace each of these parameters. In this study, the second of these parameters (the assessments' subject location) is, focused on in particular.

With the aim to contribute to the field of assessment in primary technology education, we investigate what criteria for success technology teachers identify when assessing pupils' web-based multimodal e-portfolios. The research question is:

> What criteria for success do primary teachers emphasize **during the act** of assessment?

Technology education in Sweden

In some countries, education in and about Technology is labelled *Design and Technology*, whereas in others, the terms *Technology Education* or *Engineering Education K12* are used. The designation *Technology* is the term that most closely matches teaching in what is referred to as *'teknik'* in Sweden, and will be used here**.**

Technology is a mandatory subject in Sweden from Year 1 to Year 9, i.e. compulsory school. Instruction in Technology shall provide opportunities for all pupils to develop the skills stipulated in the syllabus (NAE, 2011, pp. 256-283). The syllabus is broad, ranging through topics associated with social sciences and natural sciences as well as engineering and design. It includes theoretical and practical skills and, accordingly, instruction and assessment processes require a broad knowledge on the part of the teacher. According to school curricula, instruction should be connected to three strands of core content: *Technological solutions*, *Working methods for developing technological solutions,* and *Technology: man, society and the environment*. These core strands are divided into year spans for tutoring, Years 1–3, 4–6 and 7–9. All three strands are included in each of the three years but at different levels. The core content covers numerous topics (e.g. soldering, CAD, and man and technology). Commonly, the class teacher is responsible for teaching the subject in school Years 1-6. This obviously places great demands on individual teachers who need to acquire a broad repertoire of pedagogical and subject-specific content competencies. It should be noted that in Sweden, subjects like Art, Craft and Home Economics are independent mandatory subjects with their own syllabuses and allocated teaching time (Years 1-9). Therefore, food technology, wood, metal and textile work, for example, are not included in the Technology syllabus.

There are stipulated knowledge requirements in Years 6 and 9 (12- & 16-year-olds). The pupils' knowledge is to be holistically assessed in regard to these knowledge requirements by the teacher (Statens Skolverk (NAE), 2011). The following excerpt gives an example of the knowledge requirement for the highest grade (A) in working methods for developing technological solutions in Year 6 (12-year-olds):

> Pupils can carry out very simple work on technology and design by systematically testing and retesting possible ideas for solutions, as well as designing well-developed physical or digital models. During the work process, pupils formulate and choose action alternatives that lead to improvements. Pupils draw up well-developed documentation of the work using sketches, models or texts where the intention of the work is well documented. (NAE, 2011, p. 258).

Teachers' assessment in Technology

Harrison (2009) and Moreland, Jones and Barlex (2008) emphasise the importance of Technology teachers' own subject knowledge—not necessarily in terms of the specialist content knowledge of a subject expert but, rather, knowledge of what constitutes the subject and of pupils' misconceptions. According to research, the majority of Swedish Technology teachers lack subject-specific teacher training (Skolverket, 2013). A further complication is that Swedish Technology curricula are not well defined (Norström, 2014). Swedish primary Technology teachers are reliant upon their own prior experiences when teaching and assessing pupil's progress in the subject since support from school leaders and access to collegial discussions is scarcely provided (Hartell, 2012, 2013). Not surprisingly, Technology teachers often comment that they feel insecure while teaching the subject (Nordlander, 2011). This is troublesome because teachers' self-efficacy (i.e. their belief in their own capability to achieve) has been found to be profitable to pupils' learning (Hattie, 2009). Recent research (Hartell, Gumaelius, & Svärdh, 2014) suggests that subject-specific training in Technology has a positive effect on teachers' self-efficacy in terms of their confidence in their ability to assess their pupils' knowledge.

According to Kimbell and Stables (2008), teachers are at their most reliable when assessing pupils' work holistically, whereas they perform worst when assessing individual pieces. In relation to Technology education, Kimbell also posits that the object of assessment is closely linked to pupils' capabilities. The assessment of capabilities is complex when the focus is placed on the whole because a complete body of work is more than the sum of its constituent parts, displaying more than just knowledge, understanding or manual skills. According to Björkholm (2013), an explicit understanding of what qualities and competences constitute being 'capable in technology' is still to be explored among teachers.

Classroom assessment

Technology teachers' work methods and pedagogical approaches vary. Similarly, the documentation used in formal assessment varies in terms of both its extent and quality (Bjurulf, 2008; Hartell, 2012). According to Bjurulf, secondary school Technology teachers' assessment practices are neither transparent nor aligned with instruction methods. Even though it was found that most of the instruction provided was devoted to practical work, theoretical knowledge was valued more highly, according to the teachers' grading (Bjurulf). This focus on practical work has been made by Blomdahl (2007). In her study of Technology education in primary school (undertaken prior to the requirement for grading), the teachers studied were found to spend much time focusing on practical issues (e.g. locating material and tools), thus leaving little time for the development of tasks, assignments and assessment procedures. Problems related to Technology teachers' deficiencies in the planning and development of assessment activities have also been identified by Kimbell (2007). He contends that teachers sometimes base their assessment around whether they have taught a topic or not. Kimbell stresses the importance of careful assessment planning that embeds learning opportunities and offers consistency in assessment processes. Similarly, Harrison (2009) emphasises the importance of both planning and sharing assessment procedures with other professionals.

According to McMillan, Myran and Workman (2002), it is difficult to interpret teachers' purpose and focus when assessing pupils' knowledge and/or social progress in the classroom. They argue, as do McMillan (2005a) and Lekholm and Cliffordson (2009), that teachers use a mixture of information from different sources, sometimes blending perceptions of achievement with pupil characteristics, such as effort and participation, when grading. These inferences may, and sometimes must, be informed by factors that are both internal and external to the classroom (Bonner, 2013). Even though this blend of purposes and sources of information is necessary, it sometimes conceals the required transparency in assessment (McMillan, 2005b). Teachers' own understanding of a subject may also affect this transparency (Harrison, 2009).

The importance of teachers presenting possibilities for pupils to be active in class (to ask questions and share learning experiences, including mistakes) is emphasised in modern studies. Extensive research advocating dynamic and pupil-active teaching/learning methods has been presented by Benson (2012), Black (2008), Harrison (2009) and Moreland et al. (2008). McCormick (2004) argues for qualitative knowledge in technology education, where teachers create opportunities for pupils to think and discuss during problem-solving processes and design exercises. Benson (2012) emphasises the importance of young pupils having sufficient time to finish tasks to their own satisfaction and to reflect both individually and as a peer group. Thus, maximising the conditions for learning is of the utmost importance in any teaching and learning situation. It should be noted, however, that opportunities and tools must be available to learners to express and concretise achieved knowledge and skills.

Portfolios: A way of gathering evidence of learning in Technology Education

Teachers make numerous decisions based on inferred evidence of learning that is captured in different ways during classroom activities (Cowie, 2013; Harrison, 2009; Kimbell, 2007). Technology teachers commonly draw inferences by interpreting the winks, nods and glimpses of understanding in the eyes of their pupils (Hartell, 2013; Kimbell, 2007). These traces of decisions are important clues when studying teachers' assessment practices. However, since this kind of communication is of a subtle nature, it is very hard for researchers to capture.

Learning in technology includes developing knowledge and capabilities in terms of both processes and products (Bjurulf, 2011). Portfolios may serve as a tool for inviting learners to participate actively in the learning process; and providing authentic evidence of learning (Belgrad, 2013). Different methods for capturing and documenting both the process and product have been explored in the literature, with an emphasis on the portfolio's role in encouraging learners to be active agents in the learning process (Kimbell, 2012; Lindström, 2006; Skogh, 2008).

The e-scape model, derived from the work of The Technology Education Research Unit (TERU) at Goldsmiths College, London, UK (c.f. Kimbell, 2012), is an example. This model was developed over many years. In the beginning, different paper-based portfolios were examined. This later developed into a web-based e-portfolio system using a web-based interface. The pupils are actively involved in the documentation process, while they undertake the designated task. These portfolios are evidence-rich and are based on a designated platform that teachers, pupils and researchers can access online both within and outside school. These e-portfolios can be used for different assessment purposes, also including self- and peer assessment. The e-scape model has been used in the current study.

The inter-rated reliability among assessors of portfolios has been low, according to Belgrad (2013), and different ways of structuring portfolios and tasks have thus been investigated, with the aim of offering varied learning opportunities to all pupils. Portfolios have a tendency to be flooded with information, and Schutz and Moss (2004) suggest, therefore, that structured portfolios are preferable. Structured portfolios are an ecologically valid research instrument and a promising research tool for studying both pupils' capabilities and the assessment of pupil achievement in Technology education. Therefore, careful planning and trialling are essential in ensuring that valid tasks are provided, thus increasing the reliability of assessment processes. Tasks should provide both learning opportunities and possibilities for demonstrating accomplishments. Here, the chance to conduct valuable and reliable assessment should also be considered. Indeed, it has been found that designing assessment procedures allows teachers' understanding of learning to become clearer (Elwood, 2008).

Philosophical framework

The central research objective of this study is to gain knowledge and understanding of technology teachers' assessment of pupils' performance. Based on teachers' statements and on knowledge about the circumstances surrounding the assessment act, we aim to describe both the assessment *process* and the *outcome* of the assessment act. This approach includes an interest in teachers as individuals and in the wider situation in which teachers are situated which affects their actions.

Within the social sciences, and in the branch of social psychology normally referred to as 'symbolic interactionism' in particular, both the individual and factors that influence the individual are addressed (Mead, 1969). Ahlgren (1992) summarises the essence of symbolic interactionism:

> Individuals are who they are based on and depending on other people. Identity and consciousness are built up as a consequence of social interaction. Awareness of the self opens up for the individual to act as an active and dynamic creature, that which interprets its environment and then actively addresses it. (p. 29)

According to symbolic interactionism, actors' (in this case, teachers) understanding must be based on what they actually do: that is, on the interaction between the actor and the world. This means acknowledging the actor and the world as being engaged in dynamic processes; they are not static structures. Thus, to Mead (1969) and other proponents of symbolic interactionism, consciousness is not separated from action and interaction but, rather, is integral to both.

Although the focus of this present study lies in exploring the outcome (teachers' choice of criteria for success) of an assessment experiment rather than possible explanation for this outcome, the ideas of the symbolic interactionism have been included as an underlying framework in the design of the study. In future studies the interaction between teachers and the world in terms of the formal frames of the situation (curricula, school management and expectations from colleagues) will be further scrutinized.

Method

In the study, five teachers assessed 21 authentic multimodal pupil portfolios. In order to provide the most authentic evidence of learning possible, multimodal portfolios that capture text, drawings, voice, photos and videos were used. In the following sections, the two data collection phases are described.

Phase 1: Gathering pupils' multimodal portfolios

A teacher familiar with multimodal portfolios and with documented experience of teaching Technology was contacted. The teacher agreed to conduct the Flobot's Friend Technology task with her class of 21 pupils (Year 5, 11-year-olds). The task was selected from a collection of Technology tasks, developed, designed and evaluated by a software provider. It was also validated by the Year 5 class teacher as being an adequate task for her group of pupils; and aligned with Swedish national curricula (specifically, the goals concerning working methods for developing technical solutions). Evidence of learning was collected by the pupils, and was saved automatically on a server available to teachers, pupils and researchers via a secure login. All the pupils in the test class had prior experience of using the software (the *LiveAssess* App on iPad) and working with iPads.

In the Flobot's Friend task, pupils are asked to design and build a model of a robot friend capable of helping them in the home with particular actions. The task design encourages them to identify their own needs and to plan, document and build the model accordingly. Mind-maps, sketches, moving pictures, voice recordings and written text are used for documentation purposes, as is the model itself. There are also sub-tasks in which pupils must reflect upon and document their choices. The test class undertook the task during regular Technology lessons. In total, a sample of 21 authentic multimodal portfolios was collected for the purposes of this study.

Phase 2: Gathering evidence of teachers' criteria for success

Informants

The informants in this study are five teachers, teaching in school Years 3–6 (nine- to 12-year-olds). They all have primary-level teacher training degrees (Years 1–6). Their levels of teacher training differ to some extent; however, all of them are subject-specifically trained in Science and Mathematics. Four of the informants are also trained in Technology.[1] They each have between five and 17 years' teaching experience and work in five different schools situated across four municipalities. The teachers were *not* involved in collating the portfolios and the pupils were unknown to them (and vice versa). The results of the assessment by these teachers will not have consequences, therefore, for these pupils' present or future grades.

[1] This is very unusual. There is a well-known lack of trained Technology teachers in Sweden (see Skolinspektionen, 2014).

Adaptive Comparative Judgments (ACJ)

In order to study the teachers' evaluations during the act of assessment, the *Adaptive Comparative Judgments* method (hereafter ACJ) has been used. ACJ is described thoroughly in Pollitt (2012). It can be compared to going to the optician to get a new pair of glasses. The client is presented with different pairs of lenses until the lens that provides the best sight is found. Instead of being presented with a pair of lenses, during ACJ, the assessor/judge is presented with pairs of multimodal pupil portfolios. That is, during the assessment act, the assessor receives two portfolios at the same time. The assessor is tasked with deciding which of the two portfolios s/he considers to be the winner. Once the decision has been made, another pair of portfolios is presented. The assessment act is repeated until a certain number of pairs has been compared.

The software *LiveAssess*, provided by Digital Assess,[2] was chosen since it fulfilled our needs. The software includes both multimodalities (capturing pupil's voices, text, photos, moving pictures, mind-maps, sketches and drawings) and ACJ (Derrick, 2011). The assessors are presented with a certain number of pairs via the pair-engine; the number of pairs is determined according to the number of portfolios and assessors. The software generates statistical data for all the assessment activities. One of the products of these pair judgements is a continuum of portfolios, which resembles a rank-order but should not be used for comparing pupils' results to one another due to our assessment system. Instead, pupils' evidence of learning can be compared to the knowledge requirements of the national curricula in Technology (NAE, 2011, pp. 257–258). This act, however, is *not* within the scope of this study.

Data collection

Data collection took place in the school classroom of one of the participating teachers. In order to become acquainted with both the task and the e-portfolio format, after participating in an introductory focus group session, the five teachers undertook the Flobot's Friend task as if they were pupils themselves. After this, the rounds of ACJs began.

Each teacher sat alone, comparing two portfolios (A and B) at a time and then making a decision about which was the better (for more details, see Pollitt, 2012). For each pair, the informants were asked to think aloud regarding the motives for their decisions. They were also asked the following questions: What are you focusing on in each portfolio? What is *the* most important thing in your decision? All the statements were documented using digital sound recorders. The session lasted about 1.5 hours.

Additional ethical considerations

The study was explained to the informants, and both the ethical guidelines (including anonymity and use of collected data) and their right to withdraw from participation were made clear. Since the test-class pupils were minors, their guardians were also informed and asked to give written permission for the children to participate (this follows the ethics rules of the Swedish Academy of Sciences, see Vetenskapsrådet (2005)).

All the portfolios were made anonymous; numerical usernames were utilised instead of personal names. The actual physical models were not available for the assessors to evaluate; nevertheless, photos, videos, and pupil's verbal explanations of the models were used. The task was translated into Swedish in order to delimit possible misunderstandings due to pupils' varied abilities in English.

All the names of the informants (pupils as well as teachers) have been changed here to preserve anonymity.

---

[2] www.digitalassess.com

Analysis

Two types of data were collected: quantitative data generated by the software; and the teachers' statements during the act of assessing 21 authentic Technology portfolios.

Quantitative data

Quantitative analysis of the data provided by the software was undertaken. The judgement timing report (Table 3) provides quantitative data in terms of the number of projections; the time elapsed between pair judgements, the average time and inter-reliability data. The report shows that the inter-reliability for the ranking order was 93.2 percent. One portfolio was rated as the winner every time.

Qualitative data

The informants made 136 pair judgements altogether during 12 rounds of ACJs; and 135 think aloud responses were captured using MP3 players. These recordings (about six hours) were transcribed word-by-word. The time elapsed between each ACJ decision was noted in the transcripts. Periods of silence were also tracked, providing valuable information on decision-making. The transcripts were then transformed into *think-aloud protocols,* one for each informant.

Each teacher's ACJ judgments were numbered. This constituted the first unit of analysis. Thirteen judgments were omitted from further analysis because of software failure during the ACJ, leaving 122 (135-13) judgments for further content analysis. The recordings and transcripts of the 122 judgment acts were analysed iteratively and 159 motives were found. These motives constituted the data for further analysis.

Content analysis

Content analysis is a way of examining any form of communicative material. Content analyses may be applied to any written material and are often used to analyse large amounts of text. Presenting data in an economical way, for example in a table, often facilitates the analysis of text data (Cohen, Manion, & Morrison, 2008). This can be achieved by breaking down text into units of analysis, counting concepts and words, coding, finding patterns or clusters of similarities and differences, comparing/drawing links, and then establishing categories. Categories are the main groups of themes that represent the links between units of analysis. Cohen et al. (2008) emphasise the importance of the units fitting into the categories, as well as looking at whether a unit can be placed in more than one category. When sorting data into categories, the issue of subjectivity arises. Our categories are the result of our interpretation, which is built on our prior experience, just as the informants' interpretations and descriptions build upon their experience. In an attempt to minimise errors in interpretation, the think-aloud protocols were read by the researchers both individually and collectively in order to determine subcategories.

In qualitative research, the researcher is highly dependent on informants' willingness and ability to express their opinions and experiences verbally. An idea or experience is always a perception and it must always be studied in relation to its referent. Englund's (1993) comment "Man's view of the world is in a way more real than reality itself, for it controls their actions" (p. 70) emphasises the importance of seeking insights into how individuals perceive the reality that surrounds them.

Presenting data in another language

In order to provide illustrative examples to an international audience, translations have had to be made. The process of translating data from Swedish into English is not without problems. According to Nikander:

> Translating data extracts is not merely a question of 'adopting' or 'following' a 'transcription technique' but rather includes a range of practical and ideological questions concerning the level of detail chosen in the transcription, and of the way in which the translations are physically presented in print (2008, p. 226).

The translation to English has been kept as accurate and literal as possible; some minor linguistic adjustments have been made, however, in order to preserve the statements' tenor and style.

Results

Results from quantitative data

An analysis of the quantitative data provided by the ACJ indicates a high inter-reliability (0.932). According to the reported weighted mean square, the five teachers agree upon the rank order produced by the rounds of ACJ. Although the sample was small, none of the teachers qualified as misfits/outliers, according to the statistics presented in Table 1. Even though Evelyn was close to the boundary (1.9), her results were still within the margin of 2Sd.

Table 1: Judgement timing report providing information on the total time, projections etc.

Teacher's name	**Counted pair judgements**	**Total time hrs:min:sec**	**Average time**	**Projections**	**Weighted mean scr Z**
Annie	30	01:04:20	00:02:08	194	0.91
Evelyn	30	01:15:16	00:02:30	161	1.9
Inez	30	01:01:22	00:02:02	205	0.44
Malou	30	01:11:58	00:02:23	170	1.41
Mary	16	1:21:21	00:05:05	78	0.99
	136	05:54:17	00:02:36	808	

Not all the comparative judgments made by the teachers were captured in the recordings. Consequently, the number of pair judgments that were counted [136] does not correspond with the number of pairs analysed qualitatively [135]. Thirteen pair judgments were omitted from further analysis when software failure was put forward as the reason for choosing.

The timing report (Table 1) reveals that the teachers took their assignment seriously. What is *not* shown in Table 1 is that the more pairs the teacher was confronted with, the quicker the teacher became in making a decision. On average, the teachers took a similar amount of time to decide which of the two portfolios was the better. One teacher, Mary, needed double the assessment time compared with the rest. Mary's low-efficacy could be inferred here, but she may simply have been being thorough. Our prior experience from working with ACJ suggests that the more projections one is exposed to, the faster ACJs become. Accordingly, it is expected that Mary might have become faster if given the opportunity to practise further ACJs. This hypothesis, however, lies outside the scope of this paper.

Results from qualitative data

Qualitative analysis of the 122 pair judgements present in the think aloud protocols resulted in the establishment of 159 motives. During the first stage of content analysis, three main categories were identified: *Particular*, *Whole* and *Other*. The statements within each main category were analysed and grouped into 11 sub-categories. These sub-categories involved looking for patterns of similarities and differences and different directions of concerns, identifying the common factor within each main category. Illustrative examples of the main and sub-categories are provided in Tables 2, 3 and 4. The names of these examples follow the principle 'Main category, subcategory and number'; for instance, in Table 2, the first example from the main category is 'Particular Detail 1'.

Particular category

This first category concerns statements/motives that focus on one *particular detail or function* presented in the portfolio. This is when the teacher refers to a specific detail in a sketch or a mind-map as motive for the verdict (see illustrative example Particular Detail 1, where the pupil's ability to sketch in 3D is emphasised). There are two strands within the sub-category *function*. Illustrative examples for the sub-categories in Particular are given in Table 2. In the first strand (Function 1), the focus is placed on one particular function (a reminder function to take insulin). The second strand contains motives focusing on the number of functions (Function 2).

Table 2: Examples illustrating sub-categories present in main category 'Particular'.

Particular		Illustrative examples
Detail	1	This time I will choose Portfolio B as it include a three-dimensional sketch. However, I would have liked the student in question to develop his/her explanation regarding how it works. Here, you only get a very small and simple explanation. But the sketch outweighs it in my assessment. That is why I choose B.
Function	1	I think A wins. The reason for this is that this pupil links his project to an authentic personal problem (a reminder of when to take insulin). Otherwise, these two portfolios are quite similar in terms of models and explanations. Therefore it (A) wins.
	2	Here both sketches and mind-maps are really quite similar but Portfolio B includes … no, it is Portfolio A … also includes an additional aspect. Here, the robot can be used in two ways; as a vacuum cleaner and as a fire extinguisher. This I find quite interesting. To think outside of the box. To develop a function that could help not only the constructor him/herself, but also others. That is why I chose this one.

Whole Category

The whole category was used when the motives were not focused on one thing in particular but rather, on broader matters. Tempting, as it was to name this second category holistic, we chose not to because of the difficulty of pinpointing what the concept should include in this particular context. Based on this, we decided to name the category Whole instead of Holistic. As we intended to investigate what teachers identify without pointing them in a particular direction, we sought not to make the idea of holism explicit. This category encompasses motives relating to the process as a whole where the parts are collectively assessed as a whole unit (c.f. a puzzle) and not in isolation (as pieces of puzzle in the category Particulars). For example, the sub-category *red thread* (the pupils' work from idea to product) is considered to be present throughout the portfolio. We also allocated units framed by concepts to this category, such as *completed all sub-tasks* and *neatly done*.

The sub-category *complete* had two interpretations: firstly, completion was understood as an important virtue in and of itself; and secondly, it refers to a pupil completing the sub-tasks and thus providing more evidence of learning for the teachers to interpret. Table 3 provides illustrative examples of the sub-categories found in the whole category. This category contains the largest sample of motives (72.3 %). The *red thread* is the most frequent motive put forward by the informants when choosing one of the portfolios as a winner (53/159). *Complete* is the second most frequent motive (28/159). The complete motive can be interpreted in two ways. The first interpretation is the possibility of providing more evidence of learning for the teachers to assess. The results indicate that one reason for the teachers choosing a winning portfolio is that if a pupil has completed the task (or done more sub-tasks), more substantial evidence of learning is available to the assessor. This suggestion implies that the amount of evidence of learning is important for assessors as it guides their inferences and judgements. For completed tasks, the *red thread* (i.e. the whole process) is visible.

Table 3: Illustrative examples of sub-categories in Whole category

Whole		Illustrative examples
Red thread	1	I choose Portfolio A. A good sketch. Follows it (the sketch), and a good presentation. Detailed. The presentation as a whole is good, with a red thread present throughout the presentation.
	2	I chose Portfolio A because Portfolio B lacks a mind map, which I find important. I want to be able to see the whole process from start to finish. If you don´t present an initial idea it is difficult to see if what happens happens by coincidence or not. I want to be able to see the initial idea and how it evolves. That is why I choose Portfolio A.
	3	These are two pupils who both are able to explain the function of their robots. Still I fall for Portfolio B. The reason for this is that the pupil explains his/her decisions in different ways. It is possible to identify and follow a read thread via the given explanations, the presented description and by looking at the robot itself (the functions). So this is a documentation where you can follow his/her thoughts and intentions in relation to her construction (the robot), in writing and pictures.
Complete	1	Both A and B are well made. They include all parts. However A lack a presentation. That is why I choose B.
	2	I chose Portfolio B because it include most of the elements that facilitates a fair assessment. Portfolio A is a bit sad as only two out of the five elements have been completed.
Neat	1	This time my attention was drawn to Portfolio A. This pupil has constructed an incredibly well-made model, with creative solutions for how to use the pearls, for example, in a new and different way. So creative and also a really well-made model. This is why I chose Portfolio A.
	2	I choose Portfolio B because it is thorough and well thought out. The other one appear to be made in a rush (just to be able to finish quickly).
Explains	1	(A is chosen) Portfolio B is good too, but the explanation of what the robot is going to be used for is not clear and thorough

Other Category

Not all the units identified fitted into the two main categories and we were not to cluster these misfits together into other obvious categories, either. Consequently, an additional main category is offered, *Other.* This third category contains units that were *non-interpretable,* for example, meaning that we were unable to interpret the motives. Two illustrative examples contain motives like *more difficult to build* and *the mind-map felt more exciting*. The category was also used where there was a *dead heat* between the two portfolios, and the evaluator made an arbitrary choice. The Other category also includes motives relating to *additional values* not present in the task itself (politically correct, curricula-enforced issues, e.g. sustainability, energy and environment).

It is not always possible to explain (or know) the reasons for choosing one portfolio over another. In this study, there were four occasions where what could be called a *dead heat* occurred. Table 4 provides illustrative examples from the Other category.

Table 4: Illustrative examples of sub-categories in Other category

Other		Illustrative Examples
uninterpretable	1	I chose Portfolio A because the mind-map felt more exciting to be developed further…
	2	I chose Portfolio B. It is a good description, but I would like to have more information about how s/he made his/her robot. The material is OK, though it would have been nice if it had been painted in some colour. But that is an individual choice.
Additional	1	Portfolio B, because Portfolio B is a robot, while Portfolio A is a portable technical solution to carry with you.
	2	I chose portfolio A, given the inclusion of energy aspects and environmental issues.
	3	Portfolio A explains in more detail how the robot works, but I will still chose Portfolio B because I think that this model is much more difficult to build successfully, compared to model A. Therefore, in this case, I chose Portfolio B.
Dead Heat	1	I chose Portfolio A from spinning a coin. Head and tails. But A… well it is full of details and it has a purpose. This was hard. I do not really know how to decide. None of them 'stuck in my head', so to speak… Still, that is the sketch I picked. In fact, I really do not know what more to say.

Main and sub-categories: A summary

The number of entries for each main and sub-category, as well as for each teacher, is given in Table 5. It should be noted, however, that the purpose of this study is to clarify the areas of concern as thoroughly as possible, rather than provide quantitative information on frequency. Furthermore, the number of teachers mentioning particular types of motives has been given more importance than the number of statements included in each main category. The frequency of the motives for each category, as well as the distribution of motives provided by each teacher, is presented in Table 5.

Table 5: Frequency of sub-categories from content analysis of the 159 motives, and distribution of motives among the five teachers.

		Motives									
Main category		Whole				Particular		Other			
Teacher	Recorded judgements	*Red thread*	*Complete*	*Neat*	*Explains*	*Detail*	*Function*	*Un-interpretable*	*Additional*	*Dead heat*	Motives per teacher
Annie	26	9	5	9	1	2	1	6	3	0	36
Evelyn	27	15	0	4	3	4	3	0	0	4	33
Inez	28	13	9	0	3	4	1	1	1	0	32
Malou	25	9	13	4	2	1	2	4	3	0	38
Mary	16	7	1	6	2	2	1	0	1	0	20
Total	122	159									

Discussion

Teachers' capacity is often tacit and thus difficult to make explicit. Assessment is difficult in itself. Even the task of choosing between two portfolios has, in this study, proven to be quite a challenge. To express one's method of assessment and the qualities one values verbally is probably even more difficult.

The research question addressed in this study primarily highlights the criteria for success that primary teachers value in the act of assessment. The findings, however, call for a discussion regarding both the identified criteria for success and the potential consequences of these findings.

The 'red thread' and completing the task

The results show that the teachers in this study emphasise the importance of pupils having a *red thread* in their work. This finding is in line with Swedish curricula, being connected closely with one of the core content strands: *Working methods for developing technological solutions*. This orientation can be compared, for example, to the design processes within Design and Technology education in the UK. There are, however, some issues to address in relation to this comparison. First, the linearity within the design process could be questioned (e.g. Kimbell, 1997). Secondly, issues surrounding subject knowledge and the equipment required for pupils to develop, test and re-test technological solutions need to be highlighted. These aspects were not embedded explicitly in our test task. The lack of statements regarding the quality of the *red thread* should be considered in light of these issues. The lack of subject knowledge among Swedish Technology teachers can also be seen as a consequence of a lack of subject-specific teacher training in Technology (Skolinspektionen, 2014; Skolverket, 2013). The teachers in this study managed the situation by avoiding subject-oriented 'what-questions' and focusing on the execution of the task, whether or not the pupil followed his/her sketch, for instance, as in illustrative example *Whole 1* in Table 3; and looking at whether or not the pupils had completed the task.

Indeed, the second-most-frequent motive in the study was *having completed the task*. This criterion puts a great demand not only on the design of the task, but also on the circumstances in which the task is undertaken. In itself, *completing the task* is not a proper criterion for success; unless the task is designed very well it may be indirectly useful. Nevertheless, the results show that the teachers found it important for the pupils to finish the work. Two questions could be asked here. Firstly, are pupils given enough lesson time to complete tasks? In this particular study, some pupils, obviously, were not. Secondly, what do teachers mean by complete? Is it a quantitative or a qualitative notion? Is it a measure of neatness (or even obedience) or is it a measure of technological literacy? The answer will vary depending on an individual's views, experiences and background (Mead, 1969; Englund, 1993). This emphasis on *completing the task* points even more strongly to the importance of the *design of the task* and the *circumstances* in which tasks are undertaken.

Clarity regarding outcomes and desired abilities

Our findings indicate the need for clarity in pinpointing desired outcomes in relation to abilities. Both teachers and pupils need to be aware of these outcomes before, during, and after an assignment task. How else will teachers know what learning intentions and criteria for success to look for, as well as how to present them to their pupils so they know what is expected of them to learn and to do?

Even though the portfolios in this study were gathered under more-or-less experimental circumstances, the findings are a reminder of the importance of how Technology is taught in schools. Complete instructions (tools, materials, theoretical background, etc.) should be given to pupils in advance and during work on assignments. Equally important is the allocation of time and space for developing ideas, individually as well as with others, and allowing pupils to finish tasks to their own satisfaction.

The example *Other Additional 2* in Table 4, where the teachers added implicit criteria, is interesting. Undoubtedly, the additional criteria are connected to overarching goals for education, for instance, the teaching of environmental issues. Even though some strands of knowledge in one subject may relate to topics in other subjects, there are subject-specific matters within any subject area. Providing teaching in relation to non-subject-specific strands that arise within the teaching of a subject (for example, environmental sustainability or critical thinking in the context of Technology) is important in order to provide opportunities for pupils to learn those issues within the particular context (Wiliam, 2013).

Leaving pupils on their own while working on a task is, according to a quality report undertaken by the Swedish School Inspectorate in 2013 (Skolinspektionen, 2014), a rule rather than an exception. Based on the findings of our study, there arises the question of whether pupils actually receive instruction on how to create well-developed documentation of their work using sketches, models and text. Further, what constitutes well-developed documentation as demanded in the knowledge requirements needs to be investigated further.

Conclusions

This study covers a small number of informants, leading to limitations regarding the possibility of generalising findings. Having said this, we conclude that our findings show high reliability (0.932). It has been shown clearly that the teachers studied were unanimous in their assessment criteria. They all stressed the importance of pupils having a *red thread* in their portfolios and *completing the task*. How this should be interpreted remains open to questioning and further research.

That there seem to be common values among these teachers is encouraging. It is certainly a good foundation for the important work that remains to be done in relation to formulating assessment criteria that value not only the execution of assignments, but also subject-related skills and abilities.

The ACJ method used together with *LiveAssess* software, opens up a number of possibilities of which we have used only a small portion here. Future research will enable us to dig deeper into the mysteries of teachers' assessment processes in general and in Technology education in particular.

Affiliations

Eva Hartell
PhD
ehartell@kth.se
Department of Learning
KTH Royal Institute of Technology
Stockholm, Sweden

Inga-Britt Skogh
Professor
Department of Learning
KTH Royal Institute of Technology
Stockholm, Sweden

References

Ahlgren, RM. (1992). Jag-en social varelse. In P. Björklid & S. Fischbein (Eds.), *Individens samspel med miljön*. Stockholm, Sweden: HLS Förlag.

Belgrad, S. F. (2013). Portfolios and e-portfolios. Student reflection, self-assessment, and goal setting in the learning process. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 331–346). Thousand Oaks, CA: Sage.

Benson, C. (2012). Conformity or diversity: Developing creativity in design and technology in the early years. In H. Middleton (Ed.), *Explorations of best practice in Technology, Design & Engineering Education. Volume one* (pp. 42–51). Brisbane, Australia: Griffith Institute for Educational Research.

Bjurulf, V. (2008). *Teknikämnets gestaltningar : en studie av lärares arbete med skolämnet teknik* (1st ed.). Karlstad, Sweden. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-2729

Bjurulf, V. (2011). *Teknikdidaktik*. Riga, Latvia: Norstedts.

Björkholm, E. (2013). Exploring the capability of evaluating technical solutions: A collaborative study into the primary technology classroom. *International Journal of Technology and Design Education*, *24*(1), 1–18. doi:10.1007/s10798-013-9240-1

Black, P. (2008). Formative assessment in the learning and teaching of design and technology education: Methods and techniques. *Design and Technology Education: An International Journal*, *13*(3), 19–26.

Blomdahl, E. (2007). *Teknik i skolan: en studie av teknikundervisning för yngre skolbarn*. Stockholm: HLS förlag. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-6868

Bonner, S. M. (2013). Validity in classroom assessment: Purposes, properties, and principles. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 87–106). Thousand Oaks, CA: Sage.

Cohen, L., Manion, L., & Morrison, K. (2008). *Research methods in education* (6th ed.). Bodmin, UK: Routledge.

Cowie, B. (2013). Assessment in the science classroom. Priorities, practice, and prospects. In J. H. McMillan (Ed.), *Research on classroom assessment* (pp. 473–488). Thousand Oaks, CA: Sage.

Derrick, K. (2011). Developing the e-scape software system. *International Journal of Technology and Design Education*, *22*(2), 171–185. doi:10.1007/s10798-011-9193-1

Elwood, J. (2008). Gender issues in testing and assessment. I. In P. Murphy & K. Hall (Eds.), *Learning and practice. Agency and identities.* (pp. 87–101). London, England: Open University.

Englund, P. (1993). *Ofredsår*. Stockholm, Sweden: Atlantis.

Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (2nd ed.) (pp. 103–117). Chippenham, UK: Sage.

Harrison, C. (2009). Assessment for learning: A formative approach to classroom practice. In A. Jones & M. deVries (Eds.), *International handbook of research and development in technology education* (pp. 449–459). Rotterdam, The Netherlands: Sense.

Hartell, E. (2012). *The inefficient loneliness: A descriptive study about the complexity of assessment for learning in primary technology education.* Stockholm, Sweden. KTH Royal Institute of Technology.

Hartell, E. (2013). Looking for a glimpse in the eye: A descriptive study of teachers' work with assessment in technology education. In I.-B. Skogh & M. J. De Vries (Eds.), *Technology teachers as researchers:: Philosophical and Empirical Technology Education Studies in the Swedish TUFF Research School* (pp. 255–283). Rotterdam, The Netherlands: Sense.

Hartell, E., Gumaelius, L., & Svärdh, J. (2014). Investigating technology teachers' self-efficacy on assessment. *International Journal of Technology and Design Education*. doi:10.1007/s10798-014-9285-9

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Chippenham, Great Britain: Routledge.
Kimbell, R. (1997). *Assessing technology: International trends in curriculum and assessment.* Buckingham, England: Open University Press.
Kimbell, R. (2007). Assessment. In M. de Vries, R. Custer, J. Dakers, & G. Martin (Eds.), *Analyzing best practices in technology education* (pp. 247–258). Rotterdam, The Netherlands: Sense.
Kimbell, R. (2012). The origins and underpinning principles of e-scape. *International Journal of Technology and Design Education*, *22*(2), 123–134. doi:10.1007/s10798-011-9197-x
Kimbell, R., & Stables, K. (2008). *Researching design learning. Issues and findings from two decades of research and development*. United Kingdom: Springer.
Lekholm, A. K., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, *15*(1), 1–23. doi:10.1080/13803610802470425
Lindström, L. (2006). Creativity: What is it? Can you assess it? Can it be taught? *The International Journal of Art & Design Education*, *25*(1), 53–66.
McCormick, R. (2004). Issues of Learning and Knowledge in Technology Education. *International Journal of Technology and Design Education*, *14*(1), 21–24.
McMillan, J. H. (2005a). Secondary Teachers' Classroom Assessment and Grading Practices. *Educational Measurement: Issues and Practice*, *20*(1), 20–32. doi:10.1111/j.1745-3992.2001.tb00055.x
McMillan, J. H. (2005b). Understanding and Improving Teachers' Classroom Assessment Decision Making: Implications for Theory and Practice. *Educational Measurement: Issues and Practice*, *22*(4), 34–43. doi:10.1111/j.1745-3992.2003.tb00142.x
McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary Teachers' Classroom Assessment and Grading Practices. *The Journal of Educational Research*, *95*(4), 203–213. doi:10.1080/00220670209596593
Mead, G. H. (1969). *Mind, self and society*. Chicago, IL: University of Chicago.
Moreland, J., Jones, A., & Barlex, D. (2008). *Design and technology inside the black box Assessment for learning in the design and technology classroom*. London, England: GL Assessment.
NAE, The Swedish National Agency for Education. (2011). *Curriculum for the compulsory school, preschool class and the recreation centre, 2011* (2011). Sweden. Retrieved from http://www.skolverket.se/publikationer?id=2687 [in English]
Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, *14*(2), 149–170. doi:10.1080/09695940701478321
Nordlander, E. (2011). Vad tycker tekniklärarna? In S.-O. Hansson, E. Nordlander, & I.-B. Skogh (Eds.), *Teknikutbildning för framtiden -perspektiv på teknikundervisning i grundskolan och gymnasium.* (pp. 90–102). Stockholm, Sweden: Liber AB.
Norström, P. (2011). Technology education and non-scientific technological knowledge. *International Journal of Technology and Design Education*, 24(1), 19–38. doi:10.1007/s10798-013-9243-y
Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice*, *19*(3), 281–300. doi:10.1080/0969594X.2012.665354
Schutz, A., & Moss, P. A. (2004). Reasonable decisions in portfolio assessment: Evaluating complex evidence of teaching. *Education Policy Analysis Archives*, *12*(33), 1–47. Retrieved from http://epaa.asu.edu/epaa/v12n33/
Skogh, I.-B. (2008). Teknikdidaktisk forskning i Comenius spår. . In K. Borg & V. Lindberg (Eds.), *Kunskapande, kommunikation och bedömning i gestaltande utbildning.* (pp. 65–78). Stockholm, Sweden: Stockholms universitets förlag.
Skolinspektionen. (2014). *Teknik – gör det osynliga synligt Om kvaliteten i grundskolans teknikundervisning*. Retreived from

http://www.skolinspektionen.se/Documents/publikationssok/granskningsrapporter/kvalitetsgranskningar/2014/teknik/kvalgr-teknik-slutrapport.pdf

Skolverket. (2013). *Beskrivande data 2012. Förskola, skola och vuxenutbildning.* Retrieved from http://www.skolverket.se/publikationer?id=2994

Statens Skolverk (NAE). (2011). *Skolverkets allmänna råd Planering och genomförande av undervisning för grundskolan, grundsärskolan, specialskolan och sameskolan.* Stockholm, Sweden: Elanders.

Vetenskapsrådet. (2005). *Vad är god forskningsed? Rapport 1:2005*. Retrieved from http://www.vr.se/download/18.6b2f98a910b3e260ae28000334/god_forskningssed_3.pdf

Wiliam, D. (2013). *Principled curriculum design, Redesigning schooling-3*. London, England: SSAT (The School Network) Ltd.