



What is it called and how does it work: examining content validity and item design of teacher-made tests

Eva Hartell¹ · Greg J. Strimel²

Accepted: 12 July 2018 / Published online: 18 July 2018
© The Author(s) 2018

Abstract

This article examines content validity in teacher made tests in elementary technology education—an interdisciplinary subject mandatory for all pupils in compulsory school in Sweden. The context of teacher-based assessments relies heavily on trust for teachers to cope with demands. Even though the system is challenged and preconditions for teachers' assessment practices are not always adequate to support instruction, much is unknown about teachers' assessment practices. In this explorative study, 30 teacher-designed tests in technology education from 12 elementary schools were scrutinized in regards to content validity and the types of questions used to assess student knowledge supporting technological literacy. The results present the content validity of these tests in its current form, which may call into question the validity in terms of content and ability. Furthermore, the tests indicate how the technology school subject continues to struggle with shifting epistemologies and technologies far removed from pupils' everyday lives, which seem to contradict the aims and purpose of the subject.

Keywords Technology education · Engineering education · Test · Testing · Assessment · Content validity

Introduction

Traditionally, testing is a common way of assessing pupils' content knowledge. This field ranges from international large-scale measurement models (e.g. PISA and TIMSS) to tests created by individual teachers for use in their own classrooms. How these tests are used varies between countries, schools, and subjects. Large-scale tests can be used for comparison between schools or countries, while local tests can measure whether last week's exercises were successful or not. Although the usefulness of formal tests has been questioned (Black 2005), they are still used in most levels of the educational system. However,

✉ Eva Hartell
ehartell@kth.se

Greg J. Strimel
gstrimel@purdue.edu

¹ KTH Royal Institute of Technology, Stockholm, Sweden

² Purdue University, West Lafayette, USA

in Sweden, there are no high stakes external assessments for the school subject of technology—a mandatory subject for all Swedish pupils from year one (7 year olds) through all 9 years of compulsory schooling. Therefore, the subject's syllabus is open for various interpretations (Norström 2014) and often covers a wide variety of skills and propositional knowledge, from systematic problem solving to knowing how technology and the natural sciences have influenced each other throughout history. While the Swedish national curricula provide information about core content (ranging from the history of technology, to control and regulatory systems, constructions and types of beams, transformation of raw materials, design practices, and consequences of technologies from ecological, economic, ethical and social perspectives, such as in questions about development and use of bio-fuels and munitions). These topics (and beyond) shall be distributed across all 9 years of compulsory education, and teachers are sovereign in deciding how and when to enact a syllabus in classroom practice. This variety makes assessment and marking difficult and demands multiple ways to gather evidence of learning targeting different purposes of assessment. Unfortunately, teachers are rarely provided with training in assessment neither during teacher training nor while practicing teachers (Lundahl 2009) and access to collegial discussions, professional development, teacher training, teaching material, and support material is not offered to technology teachers (Hartell 2015). Nevertheless, the Swedish educational system trusts teachers to build a repertoire of different ways to assess student knowledge. Teachers' tacit views on learning do influence their teaching. They are more likely to articulate their points of view when it comes to learning when they are designing assessment tools (Black and Wiliam 2009; Elwood 2008; James 2010). Consequently, the purpose of this study was to examine formal teacher made tests employed in the Swedish compulsory school subject of technology in terms of construct validity.

The history of technology education in Swedish schools

Introduced in the curriculum of 1980, technology is the newest mandatory subject in Swedish compulsory school. The subject was first grouped with the science subjects such as physics, chemistry, and biology but had no proper syllabus. The implementation of the new subject was deemed by many to be unsatisfactory (Riis 1989). Teacher training was viewed as inadequate, and technology was often neglected in many schools (ASEI 2005, 2012). Technology lessons that were implemented in schools frequently focused on completing metal work or physical science related activities (Riis 1996, 2013). Many teachers lacked training in technology, which is important for developing a technology teacher's self-efficacy regarding assessment (Hartell et al. 2015).

To address these concerns surrounding the technology subject, a National Swedish Technology Curriculum was launched in 2011 (Skolverket 2011/2016). While some of the problems concerning the subject were mitigated with the national curriculum, few teachers received training on the curriculum, textbooks and tests for national assessment were unavailable, and teachers' affordance to assess pupils' knowledge was low (Hartell 2015; Skolinspektionen 2014). As a result, there are reasons to believe that the subject of technology varies more between schools than other subjects—in content as well as in complexity and how it is taught and how knowledge is assessed (Teknikdelegationen 2010; Teknikföretagen and Cetis 2013). Hartell (2015) posits that technology teachers often base their student assessments on practical work activities and documentation of this assessment is rare. She also found that teachers were left to individually manage assessment development and

implementation, which leads her to question content validity and equity in assessment from school to school and classroom to classroom.

Currently, the overarching aim of the Swedish technology subject can be described, as “helping the pupils to develop their technical expertise and technical awareness so that they can orient themselves and act in a technologically intensive world.” (Skolverket 2011/2016, p. 278). By studying the subject’s core content, pupils should develop abilities, which allow them to:

- Identify and analyse technological solutions based on their appropriateness and function,
- Identify problems and needs that can be solved by means of technology, and work out proposals for solutions,
- Use the concepts and expressions of technology,
- Assess the consequences of different technological choices for the individual, society and the environment, and
- Analyse the driving forces of technological development and how technology has changed over time. (The Swedish National Agency for Education [SNAE] 2011, p. 254)

The core content of technology can be viewed to include (but are not limited to) mechanics, materials, electronics, automatic control, technological systems, product development, and technology’s relation to the sciences, to society at large, and to the fine arts (SNAE 2011, pp. 254–264). It can be viewed as a broad and interdisciplinary subject, with no clear counterpart in higher education (Norström 2014). The syllabus of the Swedish technology subject allows for different interpretations. Most of the subject contents listed is meant to be examples (e.g., The Internet and other global technical systems, Skolverket 2011/2016, p. 281). Which these “other global technical systems” are is not decided by the Swedish National Agency for Education. Therefore, one teacher can include the railway system, another can focus on telecommunications, and a third on air traffic, and all three of them teach in accordance with the national syllabus. This gives teachers a great deal of freedom, but also a great responsibility, and both broad and deep subject knowledge is necessary in order to make informed choices.

The Swedish National Agency for Education (SNAE, *Skolverket*) has published some support material (e.g. two support booklets for the revised technology subject (Skolverket 2011, 2012), intended to aid teachers in their attempts to interpret the curriculum and apply it in a classroom context. These provide examples of suitable lesson content and guidelines for assessment. Unfortunately, these assessment guidelines do not cover whole core content in the national curriculum. Instead it concerns mainly pupils’ work with design and construction— and make suggestion on how to evaluate sketches, drawings, plans, and products, and how to provide feedback during work in these projects. Formal tests are not mentioned at all.

What is a test?

“Test” is a vague term. In school, so-called tests can be of many different kinds and intended for different purposes. In this article, the term is used in the following sense:

A test is a form of written assignment consisting of questions and/or problems to be answered/solved by pupils individually during a limited period of time. The test

takes place in a special place, such as a classroom. Either no or a limited set of tools and aids (calculator, formula sheet, ruler, ...) are allowed; normally no textbooks or notes. The purpose of the test is to assess pupils' knowledge of a certain content area, mainly on an individual basis.

This means that here we have omitted group assignments, tests that are to be answered at home with access to books and/or the Internet, and all forms of "design and make" activities that are common in technology education. Instead, only traditional tests used in classroom practice for assessing student content knowledge were examined in this study.

Testing technology in Swedish schools

Sweden has a strong tradition of classroom assessment and relies on the belief that a teacher should be given the capability to independently assess their pupil's knowledge and decide what formal grade (mark) should be awarded in accordance with regulation (Klapp Lekholm 2008, 2010; Lekholm and Cliffordson 2009). Grades are high-stake for students as grades are used for admission for further education, for upper secondary school and higher education. Since 1994, the Swedish national curriculum has been a goal- and knowledge-based criteria curriculum, in which the teachers should assess their pupils in relation to this document. According to the Swedish Educational Act (Skolfs 2010) teachers must keep track of the pupil's development in order to make it possible for the pupil to develop as far as possible within the syllabuses. The school decides when, how and by whom the pupil will be taught as long as the pupil exceeds the level of attainment stipulated in the curricula. Furthermore, the Swedish Education Act implies that the teacher is legally responsible for his/her grades, and the headmaster is responsible for their being in accordance with the law [*Skollagen* (The Swedish Education Act), chap. 3, 14–16 §§].

Generally, in Sweden, teachers are responsible both for teaching and assessment. In some subjects—mathematics, Swedish, English, and natural and social science studies—pupils partake in national tests during the spring terms of years 3, 6 and 9 (pupils are 10, 13 and 16 years old). These national tests are designed by groups of experts appointed by the SNAE, and are to provide support and monitoring for teachers' assessment practice in these subjects, but the teachers decide the final grade. There are no national tests in technology education.

In the curriculum, guidelines in terms of knowledge requirements for awarding grades are provided for every subject (SNAE 2011). The guidelines are criterion-based and there are no demands for a certain distribution among students between the different levels. In principle, all pupils could receive A (the highest grade) or all could receive F (the lowest grade, indicating that the minimum acceptable level has not yet been reached), and again it is the teachers who infer and decide. However, from statistics provided by the SNAE we can see that technology is the subject where students rarely fail, however rarely achieve highest levels of knowledge either (Hartell 2011, 2014; Skolinspektionen 2014). In technology, the criteria are related to the abilities that the subject should help the pupils develop (see above). The system is non-compensational—showing exceptional ability in one field cannot compensate for shortcomings in another, instead students are awarded grades based on lowest ability.

In technology, there is no strong tradition of testing. This is most likely to some extent because of the subject's history. It is commonly thought of primarily as a "practical" subject where skills such as design and construction are considered more important than

scholarly knowledge (Skolinspektionen 2014; Teknikföretagen and Cetus 2013). Yet, many of the themes described in the curriculum tie in well with subjects such as history, civics, and science studies, where testing-traditions are stronger (Rosenlund 2016). Testing in technology is not mentioned at all in the teachers' guides published by the SNAE (Skolverket 2011, 2012), nor in any textbooks or commonly used Swedish handbooks of technology education (Bjurulf 2011; Ginner and Mattsson 1996; Johansson and Sandström 2015). For many subjects, sample tests are commonly included in teachers' handbooks accompanying pupils' textbooks. However, in technology, the only example we have been able to find is for pupils aged 10–13 years (Sjöberg 2014). In this teachers' handbook six sample tests are provided, each covering the theme of one of the textbook's chapters: e.g. energy, traveling and transport, and written communication. The pupils' answers to each test item consist of a few sentences or a sketch; there are no multiple-choice questions. The contents are varied, and some questions ask for the pupils' opinions, which could be difficult to grade according to national curricula. No marking scheme or sample answers are provided on the information sheets accompanying the tests, but it includes references to and quotes from the curriculum, describing how the test items are linked to the subject's core contents and abilities.

Problem statement

During the last decade, there has been a rise in the interest in assessment in Sweden. This increased interest tends to focus mainly on formative assessment, and to some extent on the large-scale national and international tests and we have also seen rise in stakeholders questioning accountability of teachers' assessment practice. The significance of tests designed and used by individual teachers or in particular schools has not received much attention. Swedish teachers are supposed to design their own tests, but in spite of that, test design has traditionally not been considered an important part of teacher education (Wikström 2013) and very little research about the area exists. Still, the need for more research about teachers' assessment practices in general has been highlighted in national and international reviews of the field (Black and Wiliam 2009; Hirsh and Lindberg 2015; McMillan 2013). Within the area of technology education, the need is even greater (Bjurulf 2011; Hartell 2015; Jones et al. 2013; Ritz and Martin 2012; Williams 2011). Therefore, this study attempts to provide significant knowledge to the field.

For the Swedish technology subject, the current state of testing is unknown. In contrast to many other subjects, there are no clear model tests to be inspired by for technology teachers; mathematics teachers can and do so look at the national tests to see what an expertly designed test looks like (Boesen 2006), the French teacher has a sample test in his/her teachers' guide to be inspired from. Even the teachers of home and consumer studies have sample tests, published by the SNAE, which can serve as prototypes. Technology teachers have very little to rely on here. There are no official guidelines or examples and very little in the form of traditions—even though technology has existed as a mandatory subject for over 30 years, it is still the newest one and it has changed significantly with each new curriculum. Sample tests, such as the ones provided by Sjöberg (2014), have little impact as few schools use textbooks when teaching technology (Skolinspektionen 2014). However, again technology teachers are not obliged to use tests at all, but can base their assessment on other ways of eliciting evidence of learning and commonly do so (Hartell 2013). Even so, this study contributes to the field by exposing an important part of

teacher assessment practices particularly in technology education. While this study is centred within the Swedish education system, these assessment concerns extend beyond the country's borders as the ambiguity the school subject, internationally, continues to hinder the general understanding of the most appropriate content/practices, teacher preparation, assessment, instructional approaches, and professional development (Bartholomew et al. 2017; Strimel and Grubbs 2016; Strimel et al. 2018). Therefore, the results can be beneficial to technology educators and researchers around the world to help improve the assessment of technological literacy for all students.

Research questions

To investigate technology teacher assessments in terms of construct validity, this study used an exploratory approach and was guided by the following questions:

- What knowledge in regards to technological literacy do Swedish elementary teachers attempt to assess with self-created tests?
- What kinds of test items do the selected Swedish elementary teachers use to assess technological literacy?
- How do the teacher-made assessments in the selected Swedish Elementary Schools fulfil content validity based on the Swedish National Curricula for technology education?

Method

The data for this study were collected during a school development project titled “Boost for technology” which was aimed toward improving the quality of technology education in Sweden. To varying extents, 28 schools were involved in this project. During the project, the participating schools took part in different activities, such as completing professional development courses and participating in teacher meet-ups. Due to the strong emphasis on sovereignty for schools to decide on their own teaching and learning activities, the participating schools were asked to submit various documents regarding their instruction in technology at their schools (e.g. budget, teaching hours, local curriculums and their technology tests). These items were then used by the project's researchers to evaluate the status of technology education in Sweden.

This study focused specifically on the analysis of the testing materials submitted through the project. Tests and testing procedures can serve as a means of communication. Teachers communicate constructs of importance in assessment situations. Previous findings (e.g. Bjurulf 2008) show implicit rather than explicit practice in terms of assessment in every day classroom practice. Teacher tests are explicit in their nature, exploiting teachers' assessment practice. Also, pupils communicate information about their knowledge, skills, and abilities to the teacher. There are strong messages sent from the teacher to the pupils by assigning a test. The teacher implicitly (even if unintentionally) states that what is included in the test is important, and the pupils' ideas about what will be tested commonly influence their studies, their apprehensions about construct and the nature of the subject, and what they value. Teachers who assign tests, therefore, have a great responsibility for how this affects their pupils' attitudes and future intellectual development.

Of the 28 participating schools in the “Boost for technology” program, 16 schools responded to the request for assessment materials. The decline to respond to the call may strengthen the belief testing is marginally implemented in technology classrooms. Moreover, 3 of the 16 schools that responded stated that they did not use tests in their technology courses. The other 13 schools did provide testing materials for analysis. However, one of the school’s materials was removed from the analysis, as the submitted materials did not meet the definition of a test as described earlier. This resulted in a total of 30 tests from 12 different schools collected for analysis to address this study’s research questions. There are, however, some limitations to this study. First, the study only examines 30 tests from 12 different schools across Sweden. Second, the schools were chosen from those participating in a program aimed toward improving technology education and not randomly selected from the entire pool of Swedish schools. While the study had these limitations, the participating schools were found to be similar in terms of technology education to other Swedish schools (Hartell 2014). In addition, the sample included considerable variation in terms of socio-economic status, ethnicity, and school operation (e.g. municipality governed or privately owned schools). While these schools may serve as a diverse representation of Swedish schools (Hartell et al. 2015), the study only provides a qualitative analysis of the testing situations and therefore, it will not claim to describe typical conditions in Swedish schools in general. Still this study is of particular interest in the context of technology education internationally and particularly in current Swedish educational situation, where stricter regulations regarding who is allowed to award grades and the increased demands for teachers to document student progress and the currently growing discussion regarding equity in assessment.

To protect the participating schools in this study, the tests were coded for anonymization. The schools were sorted in a random order, and each school was assigned a name from the International Radiotelephony Spelling Alphabet: the first was called Alpha, the second Bravo, the third Charlie, etc. Each test was named after the school it originated from and a running number: Charlie-1 and Charlie-2 are two different tests from the same school, while Alpha-1 is from another, etc. Test item numbers are listed, i.e. Charlie-1:2 denotes the second question or problem in the Charlie-1 test.

All the tests were written in Swedish and have been translated for the purpose of this study. During translation, great effort has been put into avoiding problems related to different use of words in the two languages, ambiguities, etc.

Each test consists of one or more test items. The questions or problems as written on the test sheet are sometimes compounds of more than one test item. For example:

Who introduced the conveyor belt in car manufacturing facilities and what was it used for? Advantages? (Charlie-2:7)

Who invented the adjustable spanner and which country did he come from? (Echo-2:2)

In the analysis, these have been interpreted as three and two test items respectively:

- (1) Who introduced the conveyor belt in car manufacturing facilities?
- (2) What was it used for?
- (3) Which were its advantages?

- (1) What is the name of the person who invented the adjustable spanner?

(2) Which country did he come from?

This means that for many of the tests, the number of test items in the analysis is greater than the number of questions, problems, or tasks that are listed on the test sheet.

Each test item was classified according to its type and which of the technology subject's abilities in which it was related.

Classification of test items according to type

Each test item was assigned one of the following types (classification framework tool adapted from Waugh and Gronlund 2013):

- Multiple choice:** Consists of a question or problem and three or more alternatives of which one is the correct answer or solution
 Example: Which of the following gasses is heavier than air?
 Hydrogen, Marsh gas methane, Liquefied petroleum gas
 (Alpha-1:21)
- Alternative response:** Answered with one of two possible choices, often, but not always, true/false or yes/no.
 Example: Which of the following propositions is correct?
 A: A capacitor blocks alternating current but lets direct current through.
 B: A capacitor blocks direct current but lets alternating current through.
 (Hotel-1:16)
- Short answer:** Requires the respondent to supply words, numbers or symbols to answer a question or fulfil a demand.
 Examples: Name two types of bridges. (Golf-1:9)
 What does concrete consist of? (Hotel-2:11)
- Restricted-response Essay:** The response is in the form of one or more sentences, and the respondent has some freedom to decide which information to use or how to approach the problem.
 Example: Explain how a loudspeaker works. (Echo-3:3)
- Extended-response Essay:** The response is in the form of one or more sentences, and the respondent has great freedom to decide which information to use or how to approach the problem.
 Example: Originally, man used means of transportation based on his own muscle power. Later, he used animals. Describe how we have done subsequently.
 (Juliet-1:2)

Performance tasks: Solve a problem or fulfil a task.
 Example: Draw a cube with 60 mm sides using the 30/30 or 7/42 method.
 (Golf-3:1)

Other: Types of items that do not fit into the categories above but it is clear what the pupil is intended to do.

Incomprehensible: Items where the expected answer or activity is unclear to the authors of the article.

Classification using the aforementioned categories was not always easy. Especially, the boundary lines between *short answer* and *restricted-response essay*, and *restricted-response essay* and *extended-response essay*, were not always clear. These difficulties do not however affect the analyses in this article.

Classification of test items according to the technology subject's central abilities

Each test item was classified according to its connection to the following technology subject's central abilities (Skolverket 2011/2016, p. 278):

Identify and analyse technological solutions based on their appropriateness and function

(abbr.: *appropriateness and function*)

Example: Why is concrete so useful when we build houses and bridges? (Hotel-2:12)

Identify problems and needs that can be solved by means of technology, and work out proposals for solutions

(abbr.: *identify problems, work out solutions*)

Example: You are to assemble a book-case and hang two pictures on the wall. Which tools do you need? (Foxtrot-2:7)

Use the concepts and expressions of technology

(abbr.: *concepts and expressions*)

Example: What does corrugation mean? (Golf-1:5)

Assess the consequences of different technological choices for the individual, society and the environment

(abbr.: *assess consequences*)

Example: The use of hydropower has changed from the 18th century until today. Which are the positive and negative consequences for man, society, and the environment. (Kilo-1:5)

Analyse the driving forces of technological development and how technology has changed over time

(abbr.: *analyse driving forces*)

Example: Explain how the problems of bumpy roads leading to and from English coalmines were solved in the 16th century? (Delta-2:1)

Not ability-related

Items that are not ability-related may deal with individual facts that, while they are technology-related, they cannot be tied directly to any of the abilities above. Typical examples include historical facts. While they may be useful for assessing consequences and analyse driving forces, the desired response does include any assessing or analysing, instead just the repetition of facts.

Example: One of the following cities is connected to the Swedish natural gas net, which one? Gothenburg, Stockholm, Luleå (Alpha-1:25)

Not technology

Test items that deal with problems and concepts not included in the Swedish technology curriculum. Typical examples are questions about physics, chemistry and biology.

Example: Explain what the ozone layer is. (Echo-1:6d)

Vague or incomprehensible

The authors of this article have not managed to understand what the test item is about or how the response is supposed to relate to the abilities listed in the syllabus.

Example: What have I taught you that I have forgotten to ask you about? (Foxtrot-1:8)

Results

There were many similarities, but also striking differences concerning what is tested and how it is tested. Test item types ranged from simple multiple-choice questions to essays. Contents varied from facts about electrical safety to historical inventors and the physics of airplane travel. The quality of test items also varied, from questions with ambiguities and spelling mistakes, to well-written ones with clarifying examples.

The tests items and the units they consist of were also classified according to the categories described above. The results are compiled in Tables 1 and 2: one describing the types of questions used in the different tests (Table 1), and one linking question types to the abilities of the technology subject (Table 2).

Test design

The amount of time given for each test is only mentioned in a few cases. Given the extent, it seems likely that the intentions is that these are intended to be completed in one lesson. The length of technology lessons varies considerably between schools, commonly from 40 to 90 min.

The tests are intended for pupils in year 7 to year 9, but judging only from the contents it is not possible to be more specific or see progression. The Swedish syllabus (Skolverket 2011) states only what core content pupils should have the opportunity study during those 3 years, but not in which order or which school year.

Table 1 Overview of the tests

Main themes	Multiple choice	Alternative response	Short answer	Restricted response essay	Extended response essay	Performance task	Other	Incomprehensible	No of items
Alpha-1 Electricity, energy	39								39
Bravo-1 Technical drawing						3			3
Bravo-2 Domestic electricity			17	2					19
Charlie-1 Optics, acoustics			20	6	2	1			29
Charlie-2 Vehicles, transportation			18	9				2	29
Delta-1 Buildings, heating, illumination					5				5
Delta-2 Ground transportation, steam engines					5				5
Delta-3 Boats, steam engines			4		3				7
Echo-1 Energy, environmental issues			5	10				1	16
Echo-2 Innovations, solid mechanics, bridges, optics			14	2	1			4	21
Echo-3 Electromagnetism, acoustics, loudspeakers			2	1	6			1	10
Foxtrot-1 Houses and related systems			2		8			1	11
Foxtrot-2 Houses and related systems			4		7				11
Golf-1 Solid mechanics, bridges			14	3					17
Golf-2 Technical drawing						1			1
Golf-3 Technical drawing						3			3
Golf-4 Electronics			3	7	1				11
Hotel-1 Electronics		2	30	5					37
Hotel-2 Solid mechanics, materials			8	13	2			1	24
India-1 Solid mechanics			2	7					9
India-2 Electronics		4	6	3					13
India-3 Solid mechanics			12	1	1				14
Juliet-1 Transportation, vehicles					3			1	4
Juliet-2 Inventions, patents			1		2			2	5
Juliet-3 Mechanics			2		5				7

Table 1 (continued)

Main themes	Multiple choice			Short answer	Restricted response essay	Extended response essay	Performance task	Other	Incomprehensible	No of items
	Alternative response	Multiple choice	Alternative response							
Juliet-4				2	1	4				7
Kilo-1		4		6	2					12
Kilo-2				6	3	4	3			16
Lima-1				18	1		1	2		22
Lima-2				3			2	1		6
Sum	39	10		199	76	59	14	3	13	413

Table 2 Test item types and subject abilities

	Appropriateness and function	Identify problems, work out solutions	Concepts and expressions	Assess con- sequences	Driving forces and change over time	Not ability- related	Not technology	Vague	Sum
Multiple choice	6	0	8	2	0	3	20	0	39
Alternative response	10	0	0	0	0	0	0	0	10
Short answer	40	7	91	2	3	22	23	11	199
Restricted response essay	32	2	21	6	1	0	14	0	76
Extended response essay	30	3	3	3	4	0	2	14	59
Performance task	0	0	13	0	0	0	1	0	14
Other	0	2	0	0	0	0	0	2	3
Incomprehensible	3	0	1	0	2	2	1	4	13
Sum	121	14	137	13	10	27	61	30	413

The *short answer* item type was the most common (Table 1). Of the total of 413 items, 199 (48%) belonged to this category. *Short answer* and *restricted response essay*, the two categories where pupils are supposed to provide an answer in the form of one or a few words or sentences together, made up for 275 of the 413 items (67%). Almost all of these were about naming technical objects or phenomena, or providing short descriptions of their function, therefore the title of this paper “What’s it called and How does it work” helps to summarise these findings. *Extended response essays* were also mainly about technical functions, but more complex ones (e.g. methods for medical examinations using sound and radiation and possible risks associated with them [Charlie-1:16–17]). As can be seen in Table 2, approximately half of the *extended response essays* (30 out of 59) dealt with *appropriateness and function*. The two categories that are related to science, technology and society (*driving forces* and *assess consequences*) had only 4 and 3 test items respectively in the *extended response* category.

The technological domain has certain specific graphical ways of documenting and communicating, such as sketches, drawings, flow-charts, circuit diagrams, etc. These were only used to a limited extent in the tests. Also, there were a few tests that dealt with technical drawing but apart Golf-2, Golf-3, Lima-1, Lima-2, almost everything was based on text.

Lastly, none of the tests explicitly allowed any tools or aids other than common writing equipment. There was also a complete lack of “Good luck” and similar expressions of compassion from test designer to test taker.

Technological knowledge and abilities assessed in these tests

Different technological themes were seen as fit for testing (see Table 1): electricity, solid mechanics, transportation, etc. A vast majority of the test units dealt with only two of the technology subject’s abilities in relation to these themes. Of the 413 test items studied, 121 (29%) were related to the ability *appropriateness and function* (mainly describing functions), and 137 (33%) to *concepts and expressions* (mainly names and terms). Together these two categories accounted for 62% of the total number of test items. Of the 295 items that can be classified according to ability, they made up 87%. 118 of the items could not be classified to one of the abilities in technology curricula.

Analysis and discussion

The title of this article is “What is it called? How does it work?” This is chosen to highlight two common types of test items in the studied tests: to name a technological object or phenomenon (item type: *short answer*; ability: *concepts and expressions*) and to describe a function or mechanism (item type: *restricted response essay*; abilities: *appropriateness and function* and/or *concepts and expressions*).

The *extended response essays* are not as prevalent, but primarily used for functional descriptions. It is somewhat surprising that not more *extended response essays* dealt with historical aspects and political implications of technology. These themes are closely related to the subjects of history and civics, where essay-like questions are commonly used to assess pupils’ knowledge (Jansson 2012; Rosenlund 2016).

Multiple-choice items have low status in the Swedish educational system in general and they are uncommon in the national tests for compulsory school. They are used in the

university admission test (*Högskoleprovet*), which has repeatedly been criticised for this, as multiple-choice questions are said to reward fragmentary knowledge and favour men and socially privileged pupils (Andersson 2000; Högskoleverket 2000). Therefore, it is not surprising that there were few multiple-choice items in the studied tests. This can be unfortunate due to possibilities in designing multiple-choice items where each response reveals possible misunderstandings and therefore, opens possibilities for the formative use of these test.

The analysed data provided no information concerning why the different item types were chosen. Supposedly, one parameter is the type of knowledge asked for; for example, if knowing the correct name for a special kind of flying vehicle is the learning objective, there is no need for an extended essay. Another reason could be the ease of marking. Where an extended response essay may be answered in multiple ways that are all valid, the test items in the collected tests generally receive answers that are either right or wrong and it is easy to determine which is the case: the flying vehicle is called a helicopter (Delta-1:1), a square within a square on an electrical appliance indicates that it has double insulation (Bravo-2:11), a transformer is used to change the voltage (Echo-1:3). If the answer deviates from the ones stipulated in a simple template, they are wrong and not worthy of any further attention, as they do not provide evidence for/opportunity to reveal possible misconceptions. Marking is quick, comparatively objective and free from personal interpretations. Yet another possible reason is due to teachers' insecurity in determining what constitutes technological knowledge (Norström 2014) and perhaps the lack of own their subject content knowledge. Consequently, the naming of technical artefacts and describing their functions may be considered relevant knowledge in the technology subject by teachers and is hardly controversial. Other areas of knowledge, such as the history or sociology of technology can be more difficult and if can be challenging to determine when we leave the technology subject and enter the domains of others, such as history and civics.

The engineering domain is also, to a large extent, an image-based discipline. Therefore, it was somewhat surprising to see the limited use of images, illustrations and graphics in these tests. Engineers and technicians communicate using drawings, charts, and sketches. Doing this in technology education too would allow pupils to practice these skills (one of technology's *concepts and expressions* in the wording of the syllabus) while simultaneously learning about other contents. Nevertheless, these tests were dominated by text.

Test item themes

Most test items dealt with the name or the function of an individual technical artefact or artefact kind. The artefacts chosen could be referred to as traditional engineering products, such as steam engines, vehicles, electrical motors, etc. Very few dealt with modern technologies, and those that did were focused on electronics. In the studied tests, there were no test items dealing with chemical technologies, biotechnology, or computer science. The most recent invention that pupils had to show an understanding of or be able to describe the use of, was the transistor, invented in the 1940s (Golf-4:6—about the transistor as amplifier; India-2:4—description of the transistor's function). The most recent invention that was even mentioned in the tests was the integrated circuit (first patented in the late 1950s, and widely used since the 1970s), but the associated test item was only about naming the "little flat object" (microchip) (Hotel-1:2). Materials such as steel and concrete were dealt with in these tests. Modern materials, which are mentioned in the core contents for years 7–9, were nowhere to be found. The situation was the same for automatic control, pneumatics, and

computers, all of which are listed in the core contents (Skolverket 2011/2016) and absent from the tests.

Some tests included questions that are difficult to interpret and/or contained irrelevant information (so-called “window dressing”). An example of the latter was:

The first wheels were attached to a cart in Mesopotamia (Iraq) 5000 years ago. Why were rails made of wood produced?

(Charlie-2:2)

The first piece of information in the question was irrelevant for finding the answer to the question. It was also questionably conceived, as nobody can possibly know where “the first wheels” were made. “The earliest known wheels” could have been better. Wooden rails were not made until several 100 years later and how their purpose is related to the inception date for wheels is not obvious. This kind of extraneous information can be very confusing and disturbing to pupils (Haynie 1992).

The scope of the tests also varied considerably. Some consisted of test items that belonged to a certain theme (e.g. India-2 about electronics, and Juliet-3 about engineering mechanics) and were clear about what is assessed. Others seemed to be put together in a random fashion with a wide range of technological content areas being addressed. An extreme example was Echo-2, which consisted of only ten questions (21 items), but managed to include reinforced concrete, optics (binoculars and magnifying glasses), patents, and the adjustable spanner—a multitude of themes with no obvious connections.

A majority of pupils are avid users of personal computers, tablet computers, and mobile telephones. The curriculum core content could be dealt with to a large extent using examples from the ICT domain, but on these tests, it was not done. When discussing how technological development has affected society and culture, examples like the steam engine (Charlie-2:9, Delta-3:4) or reinforced concrete (Hotel-2:13) were chosen rather than mobile telephones. This could be one of the reasons why many pupils do not find the technology subject relevant to their everyday lives, which was found to be evident and expressed as great concern in two major reports by Swedish School Inspectorate (Skolinspektionen 2014) and the Government appointed Teknikdelegationen (Teknikdelegationen 2009).

Historical inventions and inventors were found on numerous tests. The items were mainly of the *short answer* type, such as “Name two Swedish inventions, and also the inventors.” (Echo-2:3) and “List as many inventions by Thomas A. Edison as possible.” (Juliet-2:2). They certainly do show some kind of knowledge about the history of technology, but it is a very fragmentary one, far from the intricacies of “Consequences of choice of technology from ecological, economic, ethical and social perspectives” and “The relationship between technological development and scientific progress” (Skolverket 2011/2016, p. 281) that are prescribed in the Swedish syllabus.

The technology tests included in this study seemingly followed the tradition of not recognizing processes, objects, and knowledge from the domestic domains of technology. There were no references to technologies related to food production, food storage, hygiene, or farming, and only a few test items related to healthcare. The view of technology purported through the collected tests was an old-fashioned one, emphasizing technologies belonging to the domains traditionally regarded as male (Berner 2009). In Sweden, the interest for technology decreases, as pupils grow older (Skolinspektionen 2014; Teknikdelegationen 2010). This is true for both boys and girls, but girls’ loss of interest is larger according to Swedish School Inspectorate (Skolinspektionen 2014). The strong focus on

“male” technologies in school is put forward as a strong contributory cause for this. The focus on “male” technologies is confirmed in this study.

Lastly, many test items were ambiguous, and in many cases, contained “window dressing” and other types of irrelevant information that could be confusing (Waugh and Gronlund 2013; Wikström 2013).

Marking

Not all tests included instructions or guidelines for evaluation and marking. In the cases where they did, it was mainly in the form of points, which are summarised for a total score. This method does not suit the Swedish criteria grading system particularly well. Each step on the official grading scale (knowledge requirements) provided by curricula implies that the pupil has knowledge of a different quality, not just a difference in quantity. To reach a higher grade, the pupil should be able to do something that is more advanced with his/her knowledge or create drawings, models etc. of a higher quality. For example: in the knowledge requirements for the end of year nine, for the grade E (lowest grade) it is stated that pupils:

can carry out simple work involving technology and design by studying and **testing** possible solutions and also designing **simple** physical or digital models (Skolverket 2011, 2016, p. 283, bold in original).

For grade A on the other hand, they have to:

carry out simple work involving technology and design by studying and **systematically testing and retesting** possible ideas for solutions and also design **well developed and well planned** physical or digital models (Skolverket 2011, 2016, p. 283, bold in original).

The difference is mainly in quality (“testing” vs. “systematically testing and retesting”, and “simple” vs. “well developed and well planned” respectively), not in quantity. This means that a sum of points (e.g. one point for knowing that copper is malleable, another for knowing that steel also has this characteristic, and a third for knowing that even aluminium can be hammered to certain shapes) cannot be mapped to the marking scale as it is intended. In some cases, teachers seemed to attempt to do this, as each test item was given a certain amount of points and the result was presented as the sum of those points. An example of this was Golf-1, where pupils were asked to describe the forces that affect the outer and inner surfaces of a bent rod (Golf-1:2). If they could conclude that the outer curve of the rod is affected by pulling forces and the inner by pushing, they would be awarded two points. In a following test item, the intersections of three types of beams were shown and students were asked to name them (Golf-1:4). Each of those correctly named (U-beam, I-beam, and tube) would be awarded with one point. Result-wise, the points of these items were interchangeable even though they showed different knowledge and belonged to different abilities (*appropriateness and function*, and *concepts and expressions*, respectively), which may question construct validity of possible inference made from that test result. Other test designers have tried to incorporate the different knowledge qualities when grading individual answers. Charlie-2 was one example, where questions about naming vehicles had E (the lowest “pass” grade) as the best possible result, while those concerning understanding could result in C, D, E, and F. This could be a way of adjusting accrediting the label E as qualitative instead of the more quantitative approach of points. However, still

the alignment with knowledge requirements E was difficult to comprehend. There were no items where it was possible to get an A or a B on that test, which may be interpreted as the teacher finding other ways to elicit higher qualitative evidence of learning.

Conclusion

The tests studied in this article played varying roles in Swedish schools, and we know that teachers infer from elicited evidence from classroom practices. Still this study has provided insights and highlights concern regarding room for improvement. For example, the design of some tests as well as the test items could be improved as these items were unclear and often relied on questions that were easy to mark rather than items that best measured the prescribed abilities and core content stipulated in the Swedish technology curricula. The following summarizes the impressions and analyses of the studied tests:

- Most tests measured only one or two of the technology subject's five abilities. When point systems were used for marking, points related to different abilities were interchangeable, which is not aligned with the intentions of the syllabus.
- The tests mainly focused on specific expressions and terms and seldom attempted to assess higher-order knowledge. Instead, pupils were often asked for specific terms or to repeat explanations of a technical artefact's functions that were directly from the textbooks. Problem solving, drawing new conclusions, evaluation, or combining information in new ways, which is demanded for the higher marks according to the grading instructions provided by national syllabus, was not encouraged in the examined test items.
- The tests did not represent the broad contents and interdisciplinary nature of the technology subject. Areas such as the sociology of technology, the relationship between technology and science, technology and gender, and large technical systems were not assessed in any of the examined tests. The reason for this could be that content is assessed in other ways. It could also indicate that it is commonly omitted when teaching the subject in the participating schools, as was the case in many schools according to Swedish School Inspectorate (Skolinspektionen 2014; Klasander 2010).
- The *concepts and expressions of technology* were in most cases limited to the knowledge of certain terms in these tests. They did not use the whole variety of concepts and expressions of technology like documentation and communication, which in many technological areas are based on images and mathematical formulas. Pupils were, therefore, not given the opportunity to practice the interpretation and creation of charts, figures, and drawings through the studied test items and the teachers would not be able to assess them with the tests they provided.
- Modern everyday technologies were absent in the studied tests. The Swedish School Inspectorate (Skolinspektionen 2014) showed that many pupils found the technology subject irrelevant and disconnected from their everyday lives. The contents of the studied tests implied several reasons for this. The technologies that pupils use every day—computers, mobile telephones, bicycles, and mopeds—were nowhere to be found in the test items. The same was true for medical technologies, sustainability, and modern materials.

Test may provide valuable insights of student learning and teachers' assessment practices. However, we must remember that tests are not the sole basis for teachers' assessment of pupils' knowledge in technology and not every part of technology is suitable to assess with tests. Teachers must therefore have an assessment repertoire suitable for purpose and use in a variety of ways to elicit evidence of learning to infer from. In many cases, tests are likely used together with the results from pupils' technical reports and "design and make" activities when grading. There is however causes for concern on behalf of construct validity and conservation of the apprehension of the subject in general and on behalf of gender bias. When one particular construct seem defined for one population and not the other, then it is better to adapt the construct instead of modifying the assessment (Wiliam 2010).

The very existence and use of these present tests might, however, (probably unintentionally) serve to conserve the apprehension of technology as a primarily "practically" oriented subject for male pupils, which is in conflict with intention of the regulation concerning compulsory schooling. A subject where the main purpose of "theoretical" knowledge is to enable the "practical" activities: the terms and expressions that are assessed in the tests are those that are mainly useful for communication in the workshop, not those used to discuss technological policy. Is this what the teacher test-designers consider to be most important? Or do they create the same kinds of tests that they have seen before, and experienced when they were pupils themselves?

To demand that pupils spend time answering questions that do not contribute to their learning in the technology subject and do not provide the teacher with any high-quality data for assessment is downright disrespectful toward the pupils and also a waste of their time. Whether the strong focus on terminology and descriptions of function provides an efficient way to teach pupils how to "orient themselves and act in a technologically advanced world" is highly questionable. To provide better opportunities for every pupil to learn what they are entitled to and receive fair grading, teachers need professional support to design appropriate learning activities and assessment procedures, including—but not limited to—tests.

Future research

Regardless of how we elicit evidence of learning in technology education, we should be more interested in how we can generalize beyond the behaviours observed during assessment activity. But still this demands adequate techniques fit for purpose. The examined tests shows there is room for improvement. However, we need not only focus on how well we assess but also on what we assess. Validity cannot be a property of a test; instead validity should be seen as a property of inferences drawn. The possible inference from these tests and relevance can be questioned and must be investigated further. These tests provided a limited view of the technology subject and provided strong evidence for construct underrepresentation. Therefore, greater focus on construct would benefit quality of the assessments themselves. Some possible inference may be valid for some purpose and for some pupils, however this study has shown too much emphasis on traditionally male technologies and construct underrepresentation in terms of modern technology present in every young person's life. Wiliam (2010) arguing the role of construct in the pursuit of equity in assessment where construct interpretation is at the heart of validity, therefore, design assessment should start with defining the construct to assess. So, a strong way forward would be to first and foremost define the construct to assess and then design an adequate method, not the other way around, as seems to be the case now in technology. Also, we

should question and improve equity and foster assessment literate teachers who can infer from evidence of learning and adapt what happens in classroom to better meet the learners needs.

This study has opened a box of opportunities for improvements. Construct of technology is somewhat defined in national curriculum, however there is room for interpretations. Perhaps focusing on a teacher's assessment repertoire could be a way to support construct definition and pupil access to fair and equitable assessment. Different constructs and purposes demand different ways of assessment; therefore, the importance of assessment literate teachers is key. Perhaps Winnie the Pooh saying "It's better to know what to look for before you start looking for it," could provide some guidance.

This is one of very few articles dealing with tests in the technology school subject, and probably the first one from the perspective of Swedish compulsory school. The authors believe that formal tests can be useful for formative as well as summative purposes in technology education, but prerequisites for this are well made tests that are aligned with the curriculum and supplemented with other types of assessment. To establish a solid foundation for how assessment in technology—including testing—should be carried out, more research is necessary. This study has dealt only with tests and test items. Further studies should consider both pupils' answers and teachers' interpretations of those answers, why these particular items were included and for what purpose were these tests used. How should or could tests contribute to pupils' learning? How should tests be designed to provide teachers with possibilities to infer relevant information about a pupil's knowledge in order to adapt instruction to better meet learners needs?

Swedish technology teachers have a multitude of backgrounds. Some of them are trained as teachers and the majority of them are not. Some combine teaching of technology with teaching of science and there are also many who combine technology and crafts, and a few who combine technology with history and civics. Therefore, it is important to ask if teachers with different backgrounds, who come from different testing traditions, design and use technology tests differently? Can their experiences from other subjects be used for the advancement of testing in technology? Perhaps most important question of all would be: How can teachers' assessment practices be supported so that every pupil prospers in technology education? How can teachers get together and help each other to build a firm foundation for assessment in technology education? If, and if so how can digital technologies be used to build assessment literate and a collective efficacious group of technology teachers?

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Andersson, P. (2000). *Att studera och bli bedömd*. Linköping: Linköping University.
- ASEI. (2005). *Alla barns rätt till teknikundervisning*. Stockholm.
- ASEI. (2012). *Teknikämnet i träda*. Retrieved from http://www.teknikforetagen.se/documents/utbildning/teknikamnet_i_trada.pdf.
- Bartholomew, S. R., Hartell, E. E., & Strimel, G. J. (2017). *ACJ: A tool for international assessment collaboration*. Paper presented at the 2017 Pupils Attitudes Towards Technology Conference, Philadelphia, PA.

- Bjurulf, V. (2008). *Teknikämnetts gestaltningar : en studie av lärares arbete med skolämnet teknik* (1st ed.). Karlstad. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-2729>.
- Bjurulf, V. (2011). *Teknikdidaktik*. Riga: Norstedts.
- Black, P. (2005). *Testing: Friend or foe? Theory and practice of assessment and testing*. New York: RoutledgeFalmer.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>.
- Boesen, J. (2006). *Assessing mathematical creativity: comparing national and teacher-made tests, explaining differences and examining impact*. Umeå universitet. Retrieved from <http://umu.diva-portal.org/smash/get/diva2:144670/FULLTEXT01.pdf>.
- 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: Open University SAGE.
- Ginner, T., & Mattsson, G. (Eds.). (1996). *Teknik i skolan*. Lund: Studentlitteratur.
- Hartell, E. (2011). Hur sätter man betyg i teknik? In S.-O. Hansson, E. Nordlander, & I.-B. Skogh (Eds.), *Teknikutbildning för framtiden- perspektiv på teknikundervisningen i grundskola och gymnasium* (pp. 75–87). Stockholm: Liber AB.
- 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* (1st ed., pp. 255–283). Rotterdam: Sense Publishers.
- Hartell, E. (2014). Exploring the (un-) usefulness of mandatory assessment documents in primary technology. *International Journal of Technology and Design Education*, 24(2), 141–161. <https://doi.org/10.1007/s10798-013-9250-z>.
- Hartell, E. (2015). *Assidere necesse est: Necessities and complexities regarding teachers' assessment practices in technology education*. KTH Royal Institute of Technology. Retrieved from <http://www.diva-portal.org/smash/get/diva2:788413/INSIDE01.pdf>.
- Hartell, E., Gumaelius, L., & Svärth, J. (2015). Investigating technology teachers' self-efficacy on assessment. *International Journal of Technology and Design Education*, 25(3), 321–337. <https://doi.org/10.1007/s10798-014-9285-9>.
- Haynie, W. J. (1992). Post hoc analysis of test items written by technology teachers. *Journal of Technology Education*, 4, 26–38.
- Hirsh, Å., & Lindberg, V. (2015). *Formativ bedömning på 2000-talet- en översikt av svensk och internationell forskning*. Stockholm, Sweden. Retrieved from <https://publikationer.vr.se/produkt/delrapport-fran-skolforskningsprojektet-formativ-bedomning-pa-2000-talet-en-oversikt-av-svensk-och-internationell-forskning/>.
- Högskoleverket [National Agency for Higher Education] (2000). *Högskoleprovet – Gårdagens mål och framtida inriktning [The university admission test: Goals of yesterday and of the future]*. Högskoleverkets rapportserie. (2000). 12 R. Högskoleverket: Stockholm.
- James, M. (2010). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 47–60). Chippenham: SAGE.
- Jansson, T. (2012). Vad kommer på provet? Gymnasielärares provpraxis i samhällskunskap. In H. Lödén (Ed.), *Forskning av och för lärare* (pp. 79–94). Karlstad: Karlstad University Press.
- Johansson, M., & Sandström, M. (2015). *Undervisa i teknik för lärare f–6*. Lund: Gleerups Utbildning.
- Jones, A., Bunting, C., & Vries, M. J. (2013). The developing field of technology education: a review to look forward. *International Journal of Technology and Design Education*, 23(2), 191–212. <https://doi.org/10.1007/s10798-011-9174-4>.
- Klapp Lekholm, A. (2008). *Grades and grade assignment: effects of student and school characteristics*.
- Klapp Lekholm, A. (2010). Vad mäter betygen? In C. Lundahl & M. Folke-Fichtelius (Eds.), *Bedömning i och av skolan- praktik, principer, politik* (pp. 129–140). Pozkal: Studentlitteratur.
- Klasander, C. (2010). *Talet om tekniska system: förväntningar, traditioner och skolverkligheter*. Linköping University Electronic Press. Retrieved from <http://liu.diva-portal.org/smash/record.jsf?pid=diva2:395176>.
- Lekholm, A. K., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15(1), 1–23. <https://doi.org/10.1080/13803610802470425>.
- Lundahl, C. (2009). *Varför nationella prov? Framväxt, dilemman, möjligheter*. Lund: Studentlitteratur AB.
- McMillan, J. H. (Ed.). (2013). Why we need research on classroom assessment. In *SAGE Handbook of Research on Classroom Assessment* (I, pp. 3–16). Thousand Oaks: SAGE Publishers.
- Norström, P. (2014). *Technological knowledge and technology education*. KTH Royal Institute of Technology. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-144875;urn:isbn:978-91-7595-078-5>.

- Riis, U. (1989). *Teknik-för alla? Tema T rapport 19*.
- Riis, U. (1996). Kan man äga ett skolämne – dragkampen om tekniken [Is it possible to own a school subject?: The tussle for technology]. In T. Ginner & G. Mattsson (Eds.), *Teknik i skolan* (pp. 41–52). Lund: Studentlitteratur.
- Riis, U. (2013). Teknik – mellan slöjd och naturvetenskap [Technology: Between sloyd and science studies]. In J. Hallström & C. Klasander (Eds.), *The Ginner Handbook of Technology Education* (pp. 100–114). Linköping University Electronic Press.
- Ritz, J. M., & Martin, G. (2012). Research needs for technology education: an international perspective. *International Journal of Technology and Design Education*, 23(3), 767–783. <https://doi.org/10.1007/s10798-012-9215-7>.
- Rosenlund, D. (2016). *History education as content, methods or orientation?: a study of curriculum prescriptions, teacher-made tasks and student strategies*. Malmö högskola.
- Sjöberg, S. (2014). *Teknik [4–6] – Lärarbok [Technology [years 4–6]: Teacher's book]*. Stockholm: Natur & Kultur.
- Skolinspektionen. (2014). *Teknik—gör det osynliga synligt Om kvaliteten i grundskolans teknikundervisning*. Stockholm.
- Skolfs. (2010). Skollagen Svensk författningssamling 2010:800 (2013). Retrieved from http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Skollag-2010800_sfs-2010-800/?bet=2010:800.
- Strimel, G. J., & Grubbs, M. E. (2016). Positioning technology and engineering education as a key force in STEM education. *Journal of Technology Education*, 27(2), 21–36.
- Strimel, G. J., Huffman, T. J., Grubbs, M. E., & Bartholomew, S. R. (2018). *Establishing progressions of learning in engineering for high school students*. Paper presented at the 2018 Pupils Attitudes Towards Technology Conference, Atholone, Ireland.
- Teknikdelegationen. (2009). *Nyfiken på naturvetenska och teknik. En kartläggning av initiativ som syftar till att öka barn och ungdomars intresse för ämnena*.
- Teknikdelegationen. (2010). *Vändpunkt Sverige—ett ökat intresse för matematik, naturvetenskap, teknik och IKT, SOU 2010:28*. Stockholm.
- Teknikföretagen & Cetis. (2013). *Teknikämnet i träda : teknikföretagens och CETIS rapport om teknikundervisningen i grundskolan*. Stockholm: Teknikföretagen.
- The Swedish National Agency for Education (SNAE). *Skolverket (2011) Curriculum for the compulsory school, preschool class and the recreation centre, 2011 (2011)*. Sweden. Retrieved from <http://www.skolverket.se/publikationer?id=2687>.
- Waugh, C. K., & Gronlund, N. E. (2013). *Assessment of student achievement* (10th ed.). Pearson.
- Wikström, C. (2013). *Konsten att göra bra prov- vad lärare behöver veta om kunskapsmätning* (1st ed.). Stockholm: Natur & Kultur.
- Wiliam, D. (2010). What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. *Review of Research in Education*, 34(1), 254–284. <https://doi.org/10.3102/0091732X09351544>.
- Williams, P. J. (2011). Research in technology education: looking back to move forward. *International Journal of Technology and Design Education*, 23, 1–9. <https://doi.org/10.1007/s10798-011-9170-8>.