

# Assessment of ICT Literacy

*A comprehensive inquiry of the educational  
readiness for the digital era*

Fazilat Siddiq



Dissertation for the degree of Philosophiae Doctor

Department of Teacher Education and School Research  
Faculty of Educational Sciences

University of Oslo

2016

© Fazilat Siddiq, 2016

*Series of dissertations submitted to the  
Faculty of Educational Sciences, University of Oslo*  
No. 261

ISSN 1501-8962

All rights reserved. No part of this publication may be  
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.  
Print production: Reprosentralen, University of Oslo.

*To mum and dad,  
my brother, Faisal,  
and my gorgeous son Aneeq.  
Thank you for all support and love,  
and for reminding me of the  
most important things in life!*



# Acknowledgements

My four-year PhD journey has come to an end and made me realize that this would not have been possible without the tremendous support and encouragement from so many people.

First, my most sincere gratitude goes to my main supervisor professor, Rolf Vegar Olsen. With a great balance between giving me the freedom and responsibility to shape my projects, and the continued support and guidance to reach my (sometimes far-fetched) goals, this has been an interesting, fun, challenging, and personally and professionally enriching journey. You have always provided academic guidance, and your support, encouragement, and constructive feedback on my work have been extremely valuable to me. I would also like to give my deepest thanks to my co-supervisor, professor Ove Edvard Hatlevik, whom I got halfway through. Your support, valuable comments, and engaging discussions have been immensely important to me. I am also grateful to my second co-supervisor, professor Ola Erstad, who introduced me to the core of the field and has offered his support and useful comments at various stages of my work.

My heartfelt thanks go to professor Mark Wilson at the Berkeley Evaluation and Assessment Research Center at the University of California, Berkeley. Thank you so much for supporting my work and generously providing resources and collaborators for all my visits to Berkeley. I feel extremely lucky to have had the chance to work on such a novel assessment and to have been guided by you. I also want to thank David Torres Irribarra for applying for the Peder Saether Center grant with me. Daniel Stansfield, thank you so much for the initial introduction to the technicalities of the assessment software.

I am infinitely grateful to my co-author and friend Ronny Scherer. You are indeed a great researcher, but for me also a role model in many ways. Thank you for increasing the quality of my work, for making me sweat and work hard, and for the many great laughs which made all the efforts worth it. Also a huge thanks goes to my co-authors Jo Tondeur, Inger Throndsen, and Perman Gochyeev for the collaboration, motivation, help, and discussions. Your support and generosity have meant a lot to me, and I feel fortunate to have had all of you as co-authors.

I need to go a little back in time and thank Svein Sjøberg, Camilla Schreiner, Ellen Henriksen, Jørgen Sjaastad, Marianne Løken, Maria Vetleseter Bøe, and several colleagues at the Norwegian Centre for Science Education for inspiring me to pursue a PhD by providing insights into the field of educational science and research in a fun and engaging way. In particular, Øystein Guttersrud, thank you for all the engaging and fun discussions and for the guidance and support. I thank you, Doris Jorde, for kicking me out of my position at the Norwegian Centre for Science and for believing that I could and should go for a PhD.

Thank you my PhD roommates at various times, Trude Nilsen, Majken Korsager, Tonje Stenseth, Stephan Daus, Lars Jenssen, Andreas Pettersen, Melaku Tesema, Nani Teig, Selina Mkimbili, and Mari Sjøberg, for the interesting conversations, motivation, and the

good laughs. Furthermore, thank you, Fredrik Jensen, for ten great years as a fellow student. I highly appreciate the many lunch walk n' talk, run for fun, or just eat n' enjoy sessions, which had little to do with research.

Sincere thanks go to four beautiful ladies for their support and care during the most stressful times. I really appreciate your generosity and friendship. Guri Nortvedt, thank you so much for all the milk for my coffee during these years. Anubha Rohatgi, those cups of tea toward the end of my work meant a lot. Jelena Radisic, thank you for providing feedback on my manuscript in the end, and Hege Kaarstein, for the help and support with managing the practicalities that come along with handing in a dissertation.

I want to thank all of the students who took the test and their teachers who showed their engagement and willingness to contribute to my research project.

Many thanks to colleagues at the Department of Teacher Education and School Research, and the Centre for Educational Measurement (CEMO) at the University of Oslo for creating a positive atmosphere, for the interesting discussions, and for the inspiration and encouragement during these years.

Also, I am very thankful to the many great researchers with a great sense of humor I have met at conferences and meetings around the world. I am forever grateful to all of you for including me so generously and making me feel a part of this growing research community.

In the end, there are people who had nothing to do with my work but still had a great influence on me and my dissertation. To all of my friends, thank you so much for the support, encouragement, and fun. In particular, thank you to my "whats-app Esssister group" and the "mountain hikers"—for helping me stay sane, content, and in touch with life outside UIO.

To my mom and dad, Safia and Siddiq; my brother, Faisal; and my sister-in law, Maryam: thank you for the never-ending support and for making me believe I could do this. The positive way of thinking, the pureness in your hearts, the cheering and guidance have shaped me in so many ways—I am truly lucky to have a family like you! Also a big thanks to the rest of my family, for the "koselig-kaos" gatherings with great food and laughs.

Thank you, Imran, for being a great friend and the best dad for our son, and for taking such good care of him so I could focus on finalizing the work.

Finally, to the little star in my life, Aneeq, my beautiful son, thank you for being the most charming, patient, and handsome little gentleman. Now, you must feel quite highly educated, first you write a master's thesis while sitting on my lap and now a PhD dissertation while sitting by my side. Thank you for making my life full of great moments and fun—you mean the world to me!

Oslo, October 2016

Fazilat Siddiq

# Abstract

This dissertation is concerned with assessment of primary and secondary students' information, communication and technology (ICT) literacy. The overarching aim of the dissertation is to investigate the positions and perspectives of different actors (teachers and students) and practices (assessment instruments) to portray how educational systems can monitor and support the development of students' ICT literacy. The background for the research focus is the importance of ICT literacy for preparing students for the digital era. Thus, the responsibility of teachers as facilitators of students' learning of ICT literacy and the critical role of assessments to monitor and seek to realize this objective is emphasized. Three individual papers contribute to the overarching aim by addressing distinct research questions and applying different methods.

The first paper systematically reviews literature on ICT literacy assessments with the aim to provide knowledge about the characteristics of the assessments, which facets of ICT literacy are measured, and the reported quality of the assessments. It draws on several theoretical frameworks and aims to bridge the disparities in the field related to the varied use of concepts and frameworks. By synthesizing research, the paper outlines the state of the art and identifies research gaps, some of which are addressed in the subsequent papers. In the second paper, an instrument to measure teachers' emphasis on the development of students' digital information and communication skills (TEDDICS) is validated. This construct describes a qualitative aspect of ICT use, and it is aligned with the ICT competences students are expected to attain (i.e., accessing, evaluating, and sharing and communicating digital information). The third paper validates the Learning in Digital Networks—ICT literacy (LDN-ICT) test, an online, performance-based assessment that measures students' ability to handle digital information, create content, communicate, and collaboratively solve problems. The findings of Paper 2 and Paper 3 revealed satisfying levels of evidence of the reliability and validity of the two scales, and further refinements and implications are suggested.

*ICT literacy frameworks* and *Assessment* emerged as two central themes across the three papers, and they form the core of the dissertation. Moreover, the Norwegian context is emphasized in the dissertation because the respondents in Paper 2 and Paper 3 are Norwegian teachers and students.

The findings of the dissertation show that the international frameworks can be aligned; theoretical and empirical evidence for the alignment is provided. Yet, in comparison, the

Norwegian ICT literacy framework has some limitations; suggestions for further revisions are given. Moreover, the importance of high-quality assessments is emphasized in the dissertation, and a set of indicators for reporting the quality of the tests was identified and further applied to appraise ICT literacy assessments. The findings show that an adequate norm for documenting and reporting the quality of ICT literacy tests is lacking. These indicators were further used as a blueprint in the two validation studies (Papers 2 and 3).

In conclusion, the dissertation contributes to the field of ICT literacy assessment by showing the interrelations between the intended, implemented, and attained curriculum, as each of them is addressed in one of the three papers. By providing state of the art in the field and validating two instruments that can be used together, the dissertation helps to inform educational systems regarding how they can monitor and support the development of students' ICT literacy.



# Table of Contents

Acknowledgements .....	IV
Abstract .....	VI
Table of Contents .....	VIII
Prologue .....	XI
1 Introduction .....	1
1.1 Background and Rationale .....	1
1.2 Delineation of the Research Field .....	5
1.3 The Three Research Papers Constituting the Dissertation .....	6
1.3.1 Overarching Aim and Research Objectives .....	6
1.3.2 A Short Summary of the Papers .....	8
1.3.3 How the Papers Collectively Address the Overarching Aim .....	10
1.4 An Outline of the Dissertation .....	12
2 Theoretical Perspectives .....	13
2.1 ICT Literacy .....	13
2.1.1 Definition of ICT Literacy .....	14
2.1.2 A Brief Historical Description of ICT Literacy .....	15
2.2 Theoretical Frameworks of ICT Literacy .....	16
2.2.1 The Revised DIGCOMP Framework .....	17
2.2.2 CIL and 21st-century Skills Frameworks .....	17
2.2.3 Comparisons among the DIGCOMP, CIL, and 21st-century skills frameworks ....	18
2.2.4 The Norwegian ICT Literacy Curriculum .....	21
2.2.5 Comparisons Between the Revised DIGCOMP Framework and the Norwegian ICT Literacy Curriculum .....	22
2.3 Assessment and Validation .....	23
2.3.1 Test Validation .....	24
3 Research Methods and Methodological Considerations .....	29
3.1 Systematic Review (Paper 1) .....	29
3.1.1 Rationale for Conducting a Systematic Review .....	29
3.1.2 Traditional Versus Systematic Reviews .....	30
3.1.3 Common Steps in Systematic Review Methodology .....	31
3.2 TEDDICS (Paper 2) .....	33

3.2.1 Exploratory Structural Equation Modeling (ESEM).....	33
3.3 LDN-ICT (Paper 3) .....	34
3.3.1 Adapting the LDN-ICT Test to the Norwegian Context.....	35
3.3.2 Questionnaire Development.....	36
3.3.3 Item Response Theory.....	37
3.4 Reflections on Research Credibility.....	38
4 Results .....	41
4.1 Summary of Main Findings.....	41
5 Discussion .....	45
5.1 An Overview of the Dissertation.....	45
5.2 ICT Literacy Frameworks .....	47
5.2.1 Alignment of the Theoretical Frameworks .....	47
5.2.2 Operationalization of ICT Literacy Frameworks .....	48
5.2.3 The Norwegian ICT Literacy Curriculum.....	51
5.3 Assessment.....	53
5.3.1 Assessment Lags Behind.....	54
5.3.2 Reporting of the Quality of Tests .....	55
5.3.3 Students' ICT Literacy .....	58
5.3.4 Teachers' Emphasis on ICT Literacy.....	60
6 Contributions, Limitations, and Closing remarks .....	62
6.1 Contributions.....	62
6.1.1 Theoretical Contribution .....	62
6.1.2 Empirical Contribution.....	63
6.1.3 Methodological Contribution .....	65
6.2 Limitations and Future Directions.....	67
6.2.1 Limitations in the Three Papers .....	67
6.3 Closing Remarks .....	69
References .....	70
APPENDICES.....	80

# Prologue

My motivation to conduct research on the assessment of ICT literacy in education grew out of a number of considerations. First, from an early age, I had the chance to “play” with a computer. This was during a time (around 1985) when computer screens were very small, hard disks were huge, and there were no graphical user interfaces. I remember reading an instruction guide in English, and somehow I learned to type “cd dir” (change directory) and other DOS commands. I also remember playing the very simple game Snake. I believe this was the starting point for my interest in computers. Second, I studied computer science at the bachelor level and pursued a master’s degree in science didactics. I have been teaching ICT at an upper secondary school in Norway. As a teacher, I have watched many of my colleagues struggle with integrating and using ICT in their classrooms—and this was not because they did not want to. Rather, they were not trained, nor did they have sufficient experience with educational technology. Finally, one of my most striking observations of ICT was related to the students I encountered. Despite the fact that they had their own laptops (or those provided by the school), in general they did not use them in a way that would enhance learning or make them more effective learners. They struggled with simple tasks such as saving files, finding the correct file, searching for information, and especially evaluating information found on the Internet. Even though students’ experience with ICT and their ICT skills varied to a large extent, this was a prevalent tendency. Hence, my observations—especially that even when students are indulged with technology, they do not necessarily have sufficient knowledge to use it in an educational context—launched my interest in ICT literacy.

It has been argued that researchers’ backgrounds, identities, values, and personal viewpoints affect their studies in several ways (Maxwell, 2013) and may be potential pitfalls for researcher bias. To understand and minimize bias, researchers should properly address and disclose their own personal stances, reflections, and expectations. Although the issue may not be as obvious in the context of a dissertation, my positive attitudes toward technology should not be confused with a naïve view on ICT as a solution for all educational concerns and issues. I believe that pedagogical use of educational technology and students’ increased ICT literacy could enhance students’ learning. However, I also very strongly believe that in certain contexts in school, it is wise to not use ICT in teaching. It is essential to find balance between

the two for the benefit of students and teachers. It is my hope this text may contribute to that discourse.

# 1 Introduction

The overarching aim of this dissertation is to investigate the assessment of information and communication technology (ICT) literacy in the educational system by examining the positions and perspectives of different actors (teachers and students) and practices (assessment instruments). This study will portray how educational systems can monitor and support the development of students' ICT literacy through relevant, purposive, and high-quality assessments.

This chapter begins with the rationale and contextual background for the dissertation, and it outlines the context in which the subsequent chapters are situated. Then, this chapter introduces the three papers discussed in the dissertation and addresses the main objectives of the dissertation. Furthermore, this chapter discusses how the three papers are related and respond to the overall aim. Finally, it presents an outline of the dissertation, including a brief overview of the chapters and how they are related.

## 1.1 Background and Rationale

ICT is pervasive in today's society and constitutes an extensive part of young peoples' lives. ICT is emphasised and formally integrated in the national curriculum of many countries (Balanskat, 2009). Hence, monitoring and assessing students' ICT literacy is critical from several viewpoints (Suto, 2013). To establish a knowledge base regarding students' actual levels of ICT literacy, it is necessary to study and summarise the characteristics of ICT literacy assessments and their underlying conceptual frameworks. Because teachers are the primary facilitators of students' ICT literacy development in the educational context, they may benefit from the information collected during these assessments. Additionally, tools must be developed to study teachers' practices and priorities with regard to their emphasis on the development of students' ICT literacy.

These perspectives (i.e., teachers, students, and assessments) are closely related to the conceptual framework of the curriculum model, which was built on the work by Goodlad, Klein, and Tye (1979). The model distinguishes between the *intended*, the *implemented*, and the *attained* curriculum (Van den Akker, 2003), which have often been applied as the domains of analysis in studies on general educational assessment (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009) and ICT in education (Law et al., 2000; Voogt & Roblin,

2012). Markauskaite (2006) emphasised that “in order to develop a comprehensive understanding of ICT literacy policies and practices in specific contexts, all three domains should be investigated” (p. 6). Voogt and Roblin (2012) made a similar recommendation when they looked at 21<sup>st</sup>-century competences and suggested that “one of the major challenges in realizing curriculum change is to ensure consistency and balance between these three curriculum representations” (p. 301). In line with this view, this PhD dissertation includes studies that span all three domains of the ICT curriculum model; these domains will be addressed more thoroughly in section 1.3.3.

In most Western societies, students tend to be *always on*, meaning that they constantly have access to communication through digital devices (Oblinger, 2004). In fact, young people in schools today have lived their lives surrounded by technology, including the Internet, computers, cell phones, tablets, smartphones, and other electronic gadgets. Considering that digital technology is ubiquitous, researchers have argued that members of the younger generation have “digital lives” (Green & Hannon, 2007), and they have been labeled as “digital natives” (Boyd, 2014). However, scholars have emphasised that being a digital native does not automatically translate into being ICT literate or digitally competent (Helsper & Eynon, 2010; Selwyn, 2009). Thus, it is important that students master ICT and develop ICT literacy to successfully participate in education, work, and society in the 21<sup>st</sup>-century (Griffin, McGaw, & Care, 2012).

It is necessary to emphasise that the pervasiveness of ICT in Western society does not apply to the entire world. There are still many countries, especially in the Third World, where the Internet is restricted to those who are wealthy and/or have a higher education (Green, 2010). The gap between people who have access to ICT resources and the Internet and those who do not have this access is called the *digital divide* (Selwyn, 2004). However, this divide has changed profoundly, especially in developed countries where access to ICT resources and the Internet are no longer substantial issues. In these countries, the concept of a digital divide has shifted from unequal access to ICT resources to unequal levels of ICT literacy (van Dijk, 2006). Students’ socio-economic backgrounds, including the language spoken at home, the number of books at home, the parents’ educational level, household income, access to ICT, and academic aspirations, have been identified as significant predictors of ICT literacy (Hatlevik & Gudmundsdottir, 2013; McLaren & Zappala, 2002; Warschauer, 2003), independent of students’ access to and frequency of ICT use. These findings stress that being inundated with technology does not inherently provide students with the competences they

need to be critical and proficient users of ICT for educational purposes (Boyd, 2014). Consequently, education has a vital and urgent role to promote ICT literacy and to prepare young people for the knowledge society.

Research has demonstrated that students lack essential skills within the ICT literacy framework. For instance, Strømsø and Bråten (2014) showed that students lacked skills related to information retrieval and processing. These findings were supported by several studies that identified gaps in students' competences in evaluating information (Aesaert, van Nijlen, Vanderlinde & van Braak, 2014; Fang, 2012), communicating and collaborating in digital environments (Calvani, Fini, Ranieri & Picci, 2012; Kuiper, Volman & Terwel, 2005).

Other aspects related to ICT literacy include the widening gap between the culture of school and the culture of students' lives outside school (Buckingham, 2007), as well as the disparities among teachers' and schools' ICT use and integration of ICT in teaching and learning (Fraillon, Ainley, Schulz, Friedman & Gebhardt, 2014). Educationalists, researchers, and policy makers are solving some of these issues by focusing on teachers' use and integration of ICT into their classroom practices (Tondeur, van Braak, Siddiq & Scherer, 2016). However, research has shown that teachers' *actual* use of ICT in their teaching and learning activities is rather limited. In a study on the integration of ICT in education in 26 countries, Pelgrum (2001) reported that teachers lacked essential ICT skills and knowledge to integrate ICT into their classroom practices. A report from the International Computer and Information Literacy study (ICILS) revealed that teachers' ICT use varied considerably; ICT was used most frequently for relatively simple tasks (e.g., word processing, presentations, information search and retrieval) and less for more complex tasks (e.g., enabling student collaboration, assessment, and feedback) (Fraillon et al., 2014).

Furthermore, because teachers are considered to be the key facilitators in promoting students' ICT literacy, several researchers have investigated factors that have hindered or promoted teachers' ICT use, as well as their attitudes and opinions related to ICT use in the classroom. In addition to resistance to change (Gomes, 2005), research has identified a lack of confidence, ICT literacy training, and teacher support as the main hindrances to the integration of ICT in classroom practices (Bingimlas, 2009; Fraillon et al., 2014). Moreover, research has shown that teachers' ICT self-efficacy and perceived usefulness of ICT are strongly related to their actual integration of ICT in learning environments (Chien, Wu, & Hsu, 2014; Scherer, Siddiq, & Teo, 2015). Ottenbreit-Leftwich, Glazewski, Newby, and Ertmer (2010) identified that teachers' ICT use was closely related to their value beliefs,

which were concerned with their own students' needs. These findings imply that teacher professional development with regard to development of teachers' technological pedagogical content knowledge (Mishra & Koehler, 2006) should not only address the development of their basic technical skills or introduction to available ICT tools but also discuss the benefits of these tools for specific purposes related to students' learning.

ICT usage in the classroom has been shown to have a positive effect on students' motivation and interest, which often results in increased attention and improved behaviour (Passey, Rogers, Machell, & McHugh, 2004). However, studies have stressed that ICT per se cannot improve learning outcomes unless it is accompanied by an underlying pedagogy (OECD, 2015; Passey et al., 2004; Smith, Higgins, Wall, & Miller, 2005; Watson, 2001).

In light of these considerations, this dissertation adopts the view that ICT literacy and use of ICT in schools require careful consideration of the various aspects which may improve its success, as well as identification of those factors that may lead to its failure. In an educational setting, success is measured in terms of indicators, such as students' actual learning outcomes, engagement, metacognition, problem solving, and critical thinking competences. It is, therefore, important that ICT is used in learning environments *only* when such learning constituents can be achieved. Thus, the educational system should ensure that all students have equal opportunities to become ICT literate by exploring, learning, and developing these competences within the educational system.

Given the dense overview of the contextual background of this study, systematic evaluations of pupils' ICT literacy in the educational system seem critical. The exact orientation of the assessments would vary based on how the *intended curriculum* is defined in the various contexts, but it is fundamental to have valid and reliable tools (e.g., assessments, teaching and learning materials and methods) to obtain the information needed at both the classroom and system levels. Hence, the focus on assessments targeting students' ICT literacy as such has become vital as they potentially serve many purposes: (1) the tests can provide insight into how the concept is operationalized in the educational system and what this means for the teachers responsible for developing students' ICT literacy; (2) how the teachers fulfill their roles as providers of the skills students should attain as part of the ICT literacy curriculum; and (3) to what extent students possess skills related to ICT, which again may inform educational policy and practice. Hence, assessments play a critical role.



## 1.2 Delineation of the Research Field

This doctoral study is concerned with a field which is broad and intersects with several other research fields. Therefore, in the following, the aim is to clarify and delineate the research area to pinpoint the focus of this study. This section should not be confused with a literature review; rather, it is a clarification of what the dissertation is concerned with and which areas of research it does not cover. A more detailed account of the literature is presented in the systematic review (Paper 1 in the dissertation) and therefore not regarded as necessary to include in this wrapping.

The subject of this study is placed in the cross-section between ICT *literacy*, *assessment*, and *education*, and each of these concepts connotes a different field. However, when integrated, they bring in different associations and meanings to different people (e.g., researchers, policy makers, educationalists, teachers). To narrow the scope of this study, a clarification of relevant and irrelevant aspects is specified. First of all, this is research on assessment of ICT literacy with reference to ICT literacy frameworks (and related concepts; see the method chapter), which is central to this study. Thus, this work is not to be concerned with research on single competence areas, such as computer-supported collaborative learning (Stahl, 2005), information literacy (Litt, 2013), use of games in education (Mislevy et al., 2016), or complex problem solving (Greiff, Wüstenberg, Molnár, Fischer, Funke, & Csapó, 2013), and neither are students' or teachers' actual classroom experiences with, perceptions of, or use of ICT. These topics are interesting and intersecting, yet out of scope of this dissertation.

There are two specific approaches to measuring ICT literacy, one which assumes that ICT is closely related to subject content (e.g., mathematics, science, reading) and another which presumes that ICT literacy transcends individual disciplines and comprises a set of knowledge, skills, and understandings that learners can adapt and transfer to new contexts (Fraillon, Schulz, & Ainley, 2013). Thus, subject content is often used as a context around the assessment items for providing students with more authentic tasks (Voogt & Roblin, 2012). In this study, the second approach is taken. ICT literacy at its core is perceived as consisting of a set of generic competences and reflects critical thinking rather than only technical or basic skills. Furthermore, ICT literacy has been assessed by both self-reports and performance-based tests. In this study, the latter assessment approach was selected, because self-reports have been shown to be biased (e.g., toward gender [Aesaert et al., 2014; Hakkarainen et al., 2000]) and may not provide an accurate picture of students' actual competence.

## 1.3 The Three Research Papers Constituting the Dissertation

The dissertation consists of three papers which aim to shed light on different aspects related to research on ICT literacy assessment in primary and secondary education. Each paper in the dissertation takes a slightly different perspective and addresses distinct research questions, discusses the results, and draws implications for the field.

The three papers included in the dissertation are:

### Paper 1

Published as

Siddiq, F., Hatlevik, O. E., Olsen, R. V., Throndsen, I., & Scherer, R. (2016). Taking a future perspective by learning from the past—A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. *Educational Research Review, 19*, 58-84. doi:10.1016/j.edurev.2016.05.002

### Paper 2

Published as

Siddiq, F., Scherer, R., & Tondeur, J. (2016). Teachers' emphasis on developing students' digital information and communication skills (TEDDICS): A new construct in 21<sup>st</sup>-century education. *Computers & Education, 92-93*, 1-14. doi:10.1016/j.compedu.2015.10.006

### Paper 3

Submitted to *Computers & Education*

Siddiq, F., Gochyyev, P., & Wilson, M. (2016). Learning in digital networks—ICT literacy: Validation of a novel assessment of 21<sup>st</sup>-century skills.

Further in the wrapping, the three papers are referred to as *systematic review*, *TEDDICS*, and *LDN-ICT*, respectively.

### 1.3.1 Overarching Aim and Research Objectives

The overarching aim is to investigate the assessment of ICT literacy by examining the positions and perspectives of different actors (teachers and students) and practices (assessment instruments) to portray a fuller picture of how educational systems can monitor and support the development of students' ICT literacy through relevant, purposive, and high-

quality assessments. Valid and reliable assessments are of vital importance for facilitating the alignment of intended, implemented, and attained curriculum. Furthermore, without high-quality assessment, it is not possible to study the complex relationship between instruction and learning of ICT literacy. Thus, the development of proper tools for assessment is of high importance for both practice and research.

More precisely, the following research objectives were posed:

Research objective 1: What is the current state of the art in the field of assessment of primary and secondary students' ICT literacy?

Research objective 2: To what extent can Teachers' Emphasis on Developing Students' Digital Information and Communication Skills (TEDDICS) be measured with high quality?

Research objective 3: To what extent can the validity of the Learning in Digital Networks–ICT literacy (LDN-ICT) test be provided?

The three research objectives addressed in the dissertation were founded in Papers 1, 2, and 3, respectively, and they collectively address the overarching aim by emphasizing the different aspects of the curriculum analysis model (Mullis et al., 2009): the intended curriculum, the implemented curriculum, and the attained curriculum as shown in Figure 1. More detailed explanations about the papers and how they are related to the levels in the model are described in section 1.3.3, and further details which explain Figure 1 will be provided, after a short summary of the three papers.

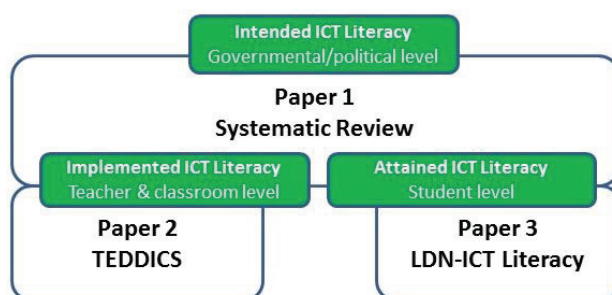


Figure 1. The interrelations among the three papers and the curriculum model, and how they address the overarching aim of the dissertation

### 1.3.2 A Short Summary of the Papers

#### *Systematic review (Paper 1)*

Paper 1 investigated the first research objective: What is the current state of the art in the field of assessment of primary and secondary students' ICT literacy? Three narrow research questions were posed to address this objective: (1) Which assessment instruments can be identified and what characterizes them? (2) which facets of ICT literacy are measured by the identified instruments? and (3) to what extent is the quality of the tests reported?

The systematic review aims to provide state-of-the-art knowledge concerning assessment of primary and secondary students' ICT literacy—focusing on an appraisal of instruments measuring these competences. An extensive search strategy was applied by combining relevant terms in a search algorithm, and it searched databases commonly used for educational research. In addition, key terms were applied to Google Web. Professional social networks such as LinkedIn, Academia.edu, and ResearchGate were also used to search for relevant articles, and discussion groups were utilized for inquiring about relevant tests for identifying *grey literature* (Jesson, Matheson, & Lacey, 2011). A set of inclusion and exclusion criteria were pre-defined and applied to the search results, and 38 tests reported in 66 studies were included in the systematic review. Data from each study were extracted by following a coding scheme, and all studies were coded by two independent researchers.

The systematic review draws on a number of theoretical considerations and serves several purposes in the dissertation. First, it has a number of theoretical contributions consisting of: (1) The revised DIGCOMP framework (Table 1), which was developed on the basis of an iterative process of categorizing the competences measured by the tests. The DIGCOMP framework was revised to enhance the clarity and applicability of the competence areas and the competences within each area. The revised DIGCOMP framework was further used to align with the following two papers and serves as the theoretical framework throughout the dissertation. (2) A coding scheme for appraising the reporting of the reliability and validity argument was developed because there was a lack of a common framework for identifying the quality across the ICT literacy assessments. (3) A four-indicator rubric for evaluating the innovativeness of the tests (i.e., design of tasks or items) was developed. The three theoretical contributions were developed on the basis of substantial literature reviews and further applied to the data (i.e., the tests) for appraising characteristics of the ICT literacy assessments. Second, the systematic review provides up-to-date knowledge about the field of ICT literacy assessment. The researchers systematically searched for and reviewed “all”

relevant literature and identified general characteristics of the studies, which facets of ICT literacy the existing tests measure, how they are measured, and the reported quality of the instruments. Gaps in the research were identified, and some of them are addressed in the two subsequent papers.

### ***TEDDICS (Paper 2)***

Paper 2 investigates the second research objective: To what extent can Teachers' Emphasis on Developing Students' Digital Information and Communication Skills (TEDDICS) be measured with high quality? In the paper, three more specific research questions were posed: (1) To what extent can the structure of the TEDDICS scale be confirmed? (2) To what extent can the external validity of TEDDICS be established? (3) Is TEDDICS invariant across teachers' gender and main subjects, and to what extent do mean differences exist?

The study draws on data from the ICILS. A representative sample of 1,100 Norwegian teachers responded to a teacher questionnaire which contained a range of variables related to teachers' ICT self-efficacy, ICT use, perceived usefulness of ICT, age, gender, and to what extent they emphasize developing their students' digital information and communications skills. The latter, which is labeled TEDDICS, was the core scale which was validated in this study. TEDDICS is a construct that describes qualitative aspects of ICT use beyond mere frequency reports, and it was conceptualized by focusing on digital skills such as accessing, evaluating, sharing, and communicating digital information (aligned with competences described in ICT literacy frameworks). The main aim of this study was to validate the TEDDICS construct, and exploratory structural equation modeling (ESEM) was applied to analyse the data. The factorial structure and generalizability of the scale were investigated, and its relations to the other constructs were examined as part of the validation process (i.e., external validity).

### ***LDN-ICT (Paper 3)***

Paper 3 addresses the third research objective in the dissertation: To what extent can the validity of the Learning in Digital Networks–ICT literacy (LDN-ICT) test be provided? Also for this study, three more targeted research questions were posed: (1) To what extent can evidence for internal validity be proved? (2) To what extent can the underlying conceptual framework be confirmed? (3) To what extent can the evidence for the external validity be provided?

Paper 3 investigates the LDN-ICT test. This study is a result of a larger collaboration with the Berkeley Evaluation and Assessment Research (BEAR) Center<sup>1</sup> at the University of California (UC), Berkeley. The original test in English was translated, adapted, and revised for use in the Norwegian language, school, and cultural setting. The LDN-ICT test comprises an online performance-based assessment in which real-time student-student collaboration is facilitated through two different platforms (CoSketch and GoogleDocs). The tasks are embedded in a test environment with open access to the Internet, and in which synchronous digital communication and collaborative problem solving are facilitated. In addition, the test attempts to measure students' ability to handle digital information and create content. A sample of 175 Norwegian students in the ninth grade took the test and responded to a questionnaire which contained background variables (e.g., gender, socio-economic status [SES]) and constructs related to their self-beliefs (collective efficacy, ICT self-efficacy, perceived usefulness of ICT, and academic aspirations). Item Response Theory (IRT) models (i.e., Unidimensional and multi-dimensional Rasch model) were applied to analyse the data. Appropriateness of the models was evaluated by examining the item fit statistics and to what extent the underlying framework consisting of four dimensions is reflected. Also, the generalizability (differential item functioning) of the measure and the relations between students' test scores and the remaining constructs were investigated as a step of inspecting the external validity of the test.

### **1.3.3 How the Papers Collectively Address the Overarching Aim**

The curriculum model is used to understand the interrelations among the three levels: the intended curriculum, the implemented curriculum, and the attained curriculum, and how the three studies in the dissertation are related to each. The model is often used by evaluation enterprises, such as the International Association for the Evaluation of Educational Achievement (IEA), to provide an in-depth understanding of students' opportunity to learn (Mullis et al., 2009). The *intended curriculum* level represents the learning goals intended for students to achieve and how the educational system should be organized to facilitate these. As illustrated in Figure 1, this view is addressed in Paper 1, which details and discusses concepts and frameworks of ICT literacy. Moreover, the DIGCOMP framework (Ferrari, 2013) was used to appraise the operationalization of the frameworks in ICT literacy assessments and

---

<sup>1</sup> For more information, visit <http://bearcenter.berkeley.edu/>

further revised to serve as a model. The next level, *implemented curriculum*, represents what is actually taught in schools or classrooms, the characteristics of those teaching it, and how it is taught. As shown in Figure 1, Paper 2, which is concerned with the degree to which teachers emphasize developing their students' ICT literacy, mainly addresses this level in the model. The teachers also responded to questions about their use of ICT in the classroom, as well as their beliefs related to their own competences and the usefulness of ICT. The TEDDICS instrument aims to capture several aspects of what is going on in the classroom with regard to ICT. In the end, the level *attained curriculum*, which is described as students' outcomes and characteristics, is included in Paper 3 (see Figure 1). Paper 3 is concerned with assessment of students' ICT literacy in digital networks. The assessment instrument aims to measure several competences within the ICT literacy framework (the intended curriculum) and may therefore inform about students' achievement related to these. The three papers aim to learn more about students' opportunities to develop ICT literacy and how the educational system could facilitate this learning. Note that the curriculum model is used to facilitate the descriptions of the linkages among the three individual papers and how they together address the overall research aim. Hence, this is a conceptual model for framing the thesis and should not be perceived as an overarching analytical model. The model is not explicitly referred to in each of the three papers but will be revisited and applied in the synthesis of the three studies, which is provided in the discussion of this wrapping.

Figure 2 provides a more detailed overview of the work presented in the dissertation. Paper 1 sets the scene for the overall dissertation by providing a comprehensive review of the research field. This work was to a large degree theoretical and concerned with synthesizing research to provide an overview of the research field, which has been fragmented because of varying use of concepts and frameworks, and identified research gaps.

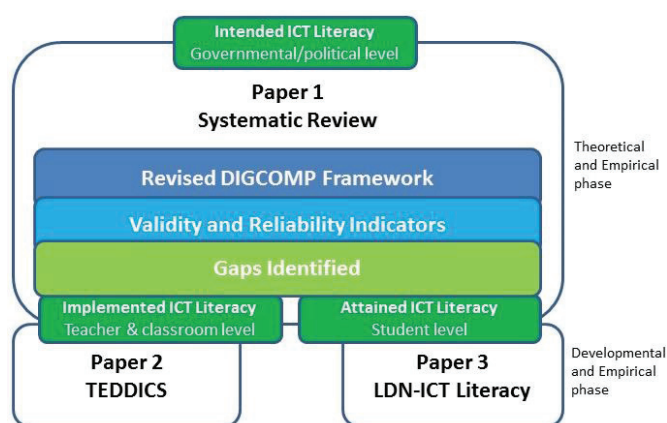


Figure 2. Dissertation overview

Note. This figure provides an overview of the dissertation, including the relationships among the three papers, the levels in the curriculum model to which they are related, and the phases they address.



Papers 2 and 3 address some of the issues highlighted in Paper 1. More specifically, they are both empirical studies, demonstrate conceptualizations of two frameworks of ICT literacy, and investigate the quality (e.g., reliability and creation of a validity argument) of the newly developed instruments. Paper 3 involved also further development of the LDN-ICT test since it was concerned with the translation, adaption, and revisions of the test as it was prepared for use with Norwegian students.

## 1.4 An Outline of the Dissertation

Chapter 1 set the scene and provided the background for this doctoral study. Moreover, the overarching research aim was stated, and an overview of how the three single studies contribute to the objectives was provided. As the field of ICT literacy is complex in terms of breadth and depth, the aim of Chapter 2 is to define and outline core terms and frameworks. In particular, the aim is to clarify the ICT literacy concept and delineate the use of this and related concepts in the dissertation. The underlying framework of the dissertation (i.e., DIGCOMP) is described, and comparisons between this and the two frameworks underlying Papers 2 and 3 are made for drawing on similarities and disparities. Finally, *assessment* is described from a theoretical perspective, and validity and reliability are discussed as crucial elements of assessment. Moreover, Chapter 2 serves as a basis for the discussion of the research findings in Chapter 5. The nature of the specific research questions, the properties of the designs, and the data available for analysis have guided the methodological choices in each of the three papers. These are described in some detail in the papers, but Chapter 3 provides a more generic presentation of them and an overarching rationale for the methodological choices. Chapter 4 provides a summary of the main results of each paper, which are further discussed in Chapter 5. Finally, in Chapter 6, the contribution of the dissertation to the field, limitations, and future directions are discussed, followed by closing remarks.



## 2 Theoretical Perspectives

In this chapter, the most relevant theories and theoretical considerations related to the dissertation are outlined. The systematic review (Paper 1) provides a thorough account of the main theoretical considerations and informs Paper 2 and Paper 3. Thus, an extensive review of the literature will not be provided. Instead, theoretical reflections regarding choice of concept, framework, and assessment are described, the purpose of which is to facilitate a broader discussion of the three studies (Chapter 5). More specifically, the *ICT literacy* concept is defined and a rationale for choosing it is given, followed by a brief historical overview of the concept and how it has evolved over the years. The DIGCOMP framework underlying Paper 1 and the overall dissertation is described, followed by a comparison with the two partly overlapping frameworks used in Paper 2 and Paper 3, the *computer and information literacy (CIL)* and *21st-century skills*, respectively. Furthermore, to facilitate a broader discussion of the relevance of this work for the national educational context, the Norwegian ICT literacy curriculum is briefly described and compared with DIGCOMP. Lastly, the theoretical perspectives concerning the field of educational assessment are outlined and delineated to the focus of this study.

### 2.1 ICT Literacy

In the research literature, myriad concepts (e.g., digital competence, ICT skills, computer literacy, ICT fluency, technological literacy, Internet skills, information literacy, media literacy) are used to describe knowledge, skills, and attitudes related to ICT (Ala-Mutka, 2011; Law, Lee, & Yuen, 2009). Even broader concepts such as new literacies, generic skills, and 21st-century skills are used to describe ICT-related competences. Moreover, efforts have been made to clarify and distinguish between the concepts in order to identify similarities and differences (Lankshear & Knobel, 2008). However, it seems to be a challenging task, and many researchers have concluded that most of the terms are used interchangeably and reflect the same content to a great extent (Law et al., 2009; Sjøby, 2013). Nevertheless, research abounds with conflicting views. Markauskaite (2006) argued that even though different terms are used synonymously for describing ICT-related capabilities, they do not necessarily convey exactly the same meaning. She further argued that the evolution of technology and society has been a driver for continuous change of the concepts and their

content. The literature also highlights the strong connection between ICT-related competences and other capabilities or literacies (e.g., numeracy, reading and science literacy, creativity, productivity, communication, collaboration skills; North Central Regional Educational Laboratory [NCREL], 2003). Thus, a change in ICT may not only affect ICT literacy, but also induce changes in more generic cognitive or non-cognitive competences. Consequently, an agreement regarding the definitions and descriptions of the terms related to ICT knowledge are seen as critical, especially in the context of education (Markauskaite, 2006) and effective policy making (Voogt & Roblin, 2012). Nevertheless, a comprehensive examination of the relevant concepts and their underlying meanings is out of scope for this dissertation.

In the dissertation, *ICT literacy* is used as the preferred term for describing students' knowledge, skills, and attitudes related to ICT in formal education. This choice was based on the following arguments: (1) ICT literacy incorporates all technologies of information processing and transmission but excludes too general or field-specific technologies (Lennon, Kirsch, Von Davier, Wagner, & Yamamoto, 2003; Markauskaite, 2006). (2) It includes the full range of knowledge, skills, abilities, attitudes, and other ICT-related capacities and is perceived as broader than only a set of ICT competences (Aesaert, 2015). (3) Other cognitive and non-cognitive attributes which may become essential due to technological development could be added to the term because it is semantically wide (Markauskaite, 2006). (4) It is seen as a life skill (such as numeracy or reading literacy) and depends on the need of the situation (Martin, 2006). (5) The term is established in the educational policy-making, decision-making, and research communities (Markauskaite, 2006).

Moreover, in this dissertation, although ICT literacy is the dominant term, digital competence and digital literacy are used synonymously, as they are closely related and connote to a large degree comparable frameworks, include converging competences, and are extensively used in educational research.

Note that the frameworks underlying the three papers in the dissertation use different concepts to address ICT literacy (e.g., digital competence, CIL, and 21st-century skills). The conceptual understanding and the content of the frameworks are described and discussed in section 2.2, which emphasizes that the concepts converge and address common domains.

### **2.1.1 Definition of ICT Literacy**

In this thesis, the following definition of ICT literacy is used: “the interest, attitude, and ability of individuals to appropriately use digital technology and communication tools to

access, manage, integrate, and evaluate information; construct new knowledge; and communicate with others in order to participate effectively in society” (Lennon et al., 2003, p. 8). This definition is in line with several other definitions of ICT literacy (Educational Testing Service [ETS], 2007; Ministerial Council on Education, Employment, Training and Youth Affairs [MCEETYA], 2005) and digital competence (Ferrari, 2013) which reflect the importance of confident and critical use of ICT for fully participating in the knowledge society. Yet, these definitions are rather general and have developed over time. Hence, a brief overview of this development will be described in next section, followed by an outline of the frameworks (i.e., DIGCOMP, CIL, and 21st-century skills) used in this dissertation to describe the content of the ICT literacy concept in more detail.

### **2.1.2 A Brief Historical Description of ICT Literacy**

Most scholars agree that operationalization of ICT literacy has changed over time due to the advancement and changes in technology (Erstad, 2006). One of the first definitions of ICT literacy was provided in the book *Digital Literacy* (Gilster, 1997). In this book digital literacy was quite generally explained as “an ability to understand and to use information from a variety of digital sources” (p. 1). Gilster was later criticized for not providing an overview of skills, competences, or attitudes to describe what it means to be digitally literate (Lankshear & Knobel, 2008). However, in his book, he discussed issues that are related to content, yet he did not systematize or further outline them.

Moreover, Gilster was not the only one of his time who was concerned with digital literacy; several other researchers used the concept (or similar concepts, see Eshet-Alkalai, 2002; Lankshear & Knobel, 2008; Martin, 2006 for an overview), defined it, and outlined the content. Nevertheless, these were in general too restrictive and too influenced by the technology of their time (Lankshear & Knobel, 2008). Thus, Gilster’s definition is more in line with the definitions of today and was supported by Eshet-Alkalai (2002), who argued that digital literacy “must be more than the ability to use digital sources effectively; it is a special kind of mindset or thinking” (p. 2). Gilster himself argued explicitly that “digital literacy is about mastering ideas, not keystrokes” (Gilster, 1997, p. 15).

Martin (2006) labeled this time period of ICT in education as *the application stage*, which lasted from the mid-1980s to the late 1990s. This stage was characterized by a focus on developing practical, basic competences in using and applying computers and software. The focus shifted from developing specialist knowledge toward using computers as everyday tools

in education, work, leisure, and home. The application stage was the successor of *the mastery stage*, which lasted from the 1960s to the mid-1980s and was characterized by schools focusing on the acquisition of knowledge about *how* computers work, as well as basic programming skills (Markauskaite, 2006; Martin, 2008). The time period since the 1990s has been labeled *the reflective stage* (Martin, 2008), which is concerned with the need for students to acquire generic skills or meta-skills to adequately cope with the changes in the knowledge society (Voogt, 2008). In other words, during this stage, the mastery of technical skills was considered insufficient with respect to developing proficient ICT literacy (ETS, 2002). Following Martin's (2008) descriptions, we are still in the *reflective stage*, labeled "late-1990s and on" (p. 157).

It has not been discussed whether this phase has been superseded by another or whether we are still in it. Albeit, many of the ideas of the reflective phase still apply today. There seems to be a new wave coming, concerned with even broader skillsets, such as 21st-century skills (Voogt & Roblin, 2012), and in which ICT literacy is often described as a subset (Griffin et al., 2012). The need for broader skillsets originated from the larger changes in society and economy due to rapid technological inventions. Moreover, these changes are further affecting work life and the types of jobs demanded in the knowledge society. As a consequence, educational systems are now challenged to prepare young people for a future with uncertain job demands (Dede, 2009).

In sum, technological advancements and the increased availability of ICT resources have created changes in learning environments. Consequently, ICT literacy and related concepts have changed, broadened, and adapted. Nevertheless, the concepts themselves do not reflect, for instance, what it is meant by being ICT literate, or which specific skills, attitudes, and competences students should attain. Thus, such concepts are most often accompanied with frameworks, which outline and detail the specific content related to the concept.

## **2.2 Theoretical Frameworks of ICT Literacy**

The three papers in the dissertation relate to different underlying frameworks. The DIGCOMP framework was chosen as the preferred framework for categorising the studies included in the systematic review (Paper 1). Paper 2 and Paper 3 were based on existing large-scale international studies which already had been defined and framed within the CIL and 21st-century skills frameworks, respectively. The following section will compare these

two frameworks with the DIGCOMP framework to ensure the comparability of the results of the three papers (Chapter 4) and to facilitate the discussion (Chapter 5). In the end, because the respondents in the two empirical studies (Paper 2 and Paper 3) were Norwegian teachers and students, the ICT literacy framework in the Norwegian curriculum is described to provide the national context in which Paper 2 and Paper 3 take place.

### **2.2.1 The Revised DIGCOMP Framework**

DIGCOMP (a Framework for Developing and Understanding Digital Competence in Europe) was initially developed by the European Commission (Ferrari, 2013). An introduction to the initial DIGCOMP framework, a justification for why the DIGCOMP framework was chosen, and, moreover, an outline of the revisions made to the original DIGCOMP framework are provided in Paper 1. Hence, in this section, the revised DIGCOMP framework is described.

The DIGCOMP framework is divided into five levels. The first level of the framework comprises six competence areas: *Information*, *Communication*, *Content Creation*, *Safety*, *Problem Solving*, and *Technical Operational Skills* (Table 1). Each competence area consists of a number of competences (level 2; Table 1), which are further fine-grained and descriptions of proficiency levels for each competence are outlined (level 3). The fourth level outlines examples of knowledge, skills, and attitudes applicable to each competence. The last and fifth level displays a contextual elaboration by providing examples of the applicability of the competence for different purposes.

DIGCOMP is a dynamic framework, which is regularly revised and updated due to rapid technological changes in the digitalization of society. A new version of DIGCOMP with initial revisions of the competence areas and competences (phase 1) was just published, and further revisions (phase 2) are planned for 2016 (Vuorikari, Punie, Carretero Gomez, & Van den Brande, 2016).

### **2.2.2 CIL and 21st-century Skills Frameworks**

The CIL framework was developed as part of the ICILS (Fraillon et al., 2013). The chosen term seems restrictive compared to other terms used more often in educational contexts. However, a rationale for the choice is provided in the framework, and the authors further argue that “ICILS was established to investigate the competences associated with

computer and information literacies as the enabling components of digital competence and 21st-century skills” (Fraillon et al., 2013, p. 16). This claim clarifies the relatedness of CIL and other broader concepts.

The top-level organizing concepts in the CIL framework are the two strands *Collecting and Managing Information* and *Producing and Exchanging Information*. Each strand is further detailed into aspects. The strand *Collecting and Managing Information* includes the aspects *Knowing about and Understanding Computer Use*, *Accessing and Evaluating Information*, and *Managing Information*. The second strand includes the aspects *Transforming Information*, *Creating Information*, *Sharing Information*, and *Using Information Safely and Securely* (Fraillon et al., 2013).

The 21st-century skills framework was developed within the international project “Assessment and Teaching of 21st-Century Skills” (ATC21S; Binkley et al., 2012), and it is based on analysis of twelve relevant frameworks drawn from a number of countries. Whereas the CIL framework aims at being very precise and closely related to the content defined as relevant for developing an assessment for secondary school students across the world, the 21st-century skills framework aims at being sufficiently broad to capture the skills anticipated to be of relevance for all citizens in the close foreseeable future. The authors argue, “Although there are significant differences in the ways in which these skills are described and clustered from one framework to another, we consider that the above list of ten is sufficiently broad and comprehensive to accommodate all approaches” (Binkley et al., 2012, p. 36). The framework distinguishes between four categories containing ten skillsets: Ways of Thinking (including the skillsets *Creativity and Innovation*; *Critical Thinking*, *Problem Solving*, and *Decision Making*; *Learning to Learn*; and *Metacognition*); Ways of Working (including the skillsets *Communication* and *Collaboration*); Tools for Working (including the skillsets *Information Literacy* and *ICT Literacy*); Living in the World (including the skillsets *Citizenship*, *Life and Career*, and *Personal and Social Responsibility*) (Binkley et al., 2012, p. 18). Each skillset is further detailed with regard to competences, skills, attitudes, and values.

### **2.2.3 Comparisons among the DIGCOMP, CIL, and 21st-century skills frameworks**

Although the three frameworks were developed within different larger projects for rather different purposes, they convey many of the same principles, are all related to



education, and, moreover, aim to describe *what* and *how*, students acquire, use, adapt to, and learn with technology.

The revised DIGCOMP framework (Ferrari, 2013) with the six competence areas is the starting point for a brief comparison, as shown in Table 1. It is beyond the scope of this thesis to provide an exhaustive comparison of the three frameworks with a full review and synthesis of every detail included in them. The following overview's sole purpose is to provide enough detail in order to highlight the similarities and differences of specific relevance to allow for an overarching discussion of the findings in the three papers in the thesis.

In general, the three frameworks emphasise students' ability to collect and understand information; to produce information; to communicate digital information; and to search, produce, and communicate in a safe and responsible way. The competence area *Information* involves identifying, retrieving, and analysing digital information (Ferrari, 2013). This competence area is equally covered in the three frameworks (Table 1). The competence area *Communication* refers to students' awareness, knowledge, and understanding of communication with others. Similar descriptions can be found in the CIL framework aspect *Sharing Information* (2.3) and in the 21st-century skills framework aspects *Communication* (4) and *Collaboration* (5). Note that the collaboration aspect of communication is lacking in the CIL framework. The third competence area, *Content Creation*, captures the students' use of digital tools for production, publishing, and problem solving. Similar descriptions could be found in the CIL framework aspects *Transforming Information* (2.1) and *Creating Information* (2.2), and also in the 21st-century skills framework *Creativity and Innovation* (1) together with *Critical Thinking* (2). The area *Safety* involves personal protection, data protection, digital identity, and security issues (Ferrari, 2013). Similar descriptions can be found in the CIL framework aspect *Using Information Safely and Securely* (2.4) and as part of the skills in Personal and Social Responsibility (10) (Binkley et al., 2012). Furthermore, *Problem Solving* involves the ability to identify and solve various problems. Similar descriptions can be found in the 21st-century skills framework category 2, but not in the CIL framework. Finally, for the area *Technical Operational Skills*, similar descriptions can be found in the CIL framework (e.g., *Knowing about and Understanding Computer Use* [1.1]), but this is not explicitly mentioned in the 21st-century skills framework.

Table 1. A comparison among the revised DIGCOMP framework and CIL, 21st-century skills, and the Norwegian ICT literacy curriculum

DIGCOMP	CIL	21 <sup>st</sup> Century skills	Norwegian curriculum
Competence areas and competences	Strands and Aspects	Categories and skillsets	Categories
<b><u>1. Information</u></b> 1.1 Browsing, searching, and filtering information 1.2 Evaluating information 1.3 Storing and retrieving information	<b><u>Collecting and managing information</u></b> 1.1 Knowing about and understanding computer use 1.2 Accessing and evaluating information 1.3 Managing information	<b><u>Tools for Working</u></b> 6. Information literacy 7. ICT literacy	<b><u>Search and process</u></b>
<b><u>2. Communication</u></b> 2.1 Interacting through digital technologies 2.2 Sharing information and content 2.3 Engaging in online citizenship 2.4 Collaborating through digital technologies *2.1.1 Asynchronous communication *2.1.2 Synchronous communication *2.4.1 Asynchronous collaboration *2.4.2 Synchronous collaboration	2.3 Sharing information	<b><u>Ways of Working</u></b> 4. Communication 5. Collaboration	<b><u>Communicate</u></b>
<b><u>3. Content creation</u></b> 3.1 Developing content 3.2 Integrating and re-elaborating 3.3 Copyright and licenses 3.4 Programming	<b><u>Producing and exchanging information</u></b> 2.1 Transforming information 2.2 Creating information	<b><u>Ways of Thinking</u></b> 1. Creativity and innovation 2. Critical thinking, problem solving, decision making	<b><u>Produce</u></b>
<b><u>4. Safety</u></b> 4.1 Protecting devices <b>4.2 Managing and protecting personal data</b> 4.3 Protecting health 4.4 Protecting the environment <b>4.5 Netiquette</b>	2.4 Using information safely and securely	<b><u>Living in the World</u></b> 10. Personal and social responsibility	<b><u>Digital judgment</u></b>
<b><u>5. Problem solving</u></b> <b>5.1 Solving problems with use of digital technology</b> <b>5.2 Collaborative problem solving</b> 5.3 Innovating and creatively using technology 5.4 Identifying digital competence gaps		2. Critical thinking, problem solving, decision making	
<b><u>6. Technical operational</u></b> <b>6.1 Solving technical problems</b> <b>6.2 Identifying needs and technological responses</b> <b>6.3 Basic technical skills</b>	1.1 Knowing about and understanding computer use		

*Note.* The competences in bold letters represent the revisions of the original DIGCOMP framework. Elements marked with an asterisk (\*) refer to level 3 in DIGCOMP.



## 2.2.4 The Norwegian ICT Literacy Curriculum

ICT literacy has been integrated in the national curriculum in Norway since the educational reform labeled as knowledge promotion in 2006. Norway was one of the first countries to enhance the status of ICT literacy by including it in the national curriculum (Balanskat & Gertsch, 2010; Krumsvik, 2008), and it is the teachers who are formally responsible for teaching ICT literacy. ICT literacy is not a subject in the compulsory education, but it is defined as one of the basic key literacies (along with reading, writing, oral skills, and numeracy) to be integrated with the competence aims of the school subjects (Norwegian Directorate for Education and Training, 2012). This represents a strong signal about the content and direction for the future development of ICT as key literacy for learning (Erstad, 2010).

In the national curriculum, ICT literacy is defined as follows: “Digital skills involve being able to use digital tools, media, and resources efficiently and responsibly, to solve practical tasks, find and process information, design digital products, and communicate content. Digital skills also include developing digital judgment by acquiring knowledge and good strategies for the use of the Internet” (Norwegian Directorate for Education and Training, 2012, p. 12). Moreover, ICT literacy is described as a prerequisite for further learning and for active participation in working life and society.

The ICT literacy framework is outlined in a grid and consists of four categories, namely: *Search and Process*, *Produce*, *Communicate*, and *Digital Judgment*. For each of these categories, descriptions of the progression through five levels are provided, and each competence category (i.e., the cells) in the grid formulates performance standards at that level (Appendix A). The framework continues by stating that “the requirements are general and serve as a basis and point of reference for developing subject and grade relevant competence aims” (Norwegian Directorate for Education and Training, 2012, p. 5). Moreover, it is stated in the framework that each subject curriculum group needs to “make decisions on which grids, cells, and levels are relevant for their subject as well as for different age groups of students, and formulate competence aims based on these decisions” (Norwegian Directorate for Education and Training, 2012, p. 5).

## 2.2.5 Comparisons Between the Revised DIGCOMP Framework and the Norwegian ICT Literacy Curriculum

A comparison between the revised DIGCOMP framework and the Norwegian ICT literacy curriculum (Table 1) indicates that the competence areas *Information*, *Content Creation*, and *Safety* in the DIGCOMP framework are to some extent covered in the Norwegian ICT literacy curriculum, whereas the competence areas *Communication* and *Problem Solving* are to a great extent lacking. Hence, the Norwegian ICT literacy framework lacks descriptors related to the more generic 21st-century skills (e.g., *Communication*, *Collaboration*, *Problem Solving*, *Creativity*; competence areas 4 and 5 in the revised DIGCOMP framework).

The largest difference between the two is that the Norwegian curriculum is less detailed, and many of the single competences in DIGCOMP are included in the level descriptions instead (see Appendix A for an overview of the Norwegian ICT literacy framework). For instance, the category *Search and Process* is described at level 1 as “can read hypertexts and simpler interactive information ...,” whereas the description at level 2 is “can make simple digital searches, and read and interpret information from digital sources ...” and at level 3 continues by stating “can choose and use search strategies and assess information from digital sources ... .” Similarly, levels 4 and 5 describe further ability expectations of students at these levels. While *Browsing, Searching and Filtering Information* (1.1) and *Evaluating Information* (1.2) are formulated as two separate competences in DIGCOMP, they are regarded as levels of higher complexity under the content category *Search and Process* in the Norwegian curriculum. This indicates that students at higher ICT literacy levels are expected to manage evaluation of information and use proper search strategies, whereas students at lower levels are not expected to be able to search for or assess information. These findings point toward the insufficient structure of the Norwegian framework, which challenges further comparisons. In particular, one categorisation level corresponding to *competences* in DIGCOMP, *aspects* in CIL, or *skillsets* in the 21st-century skills framework is missing (see Table 1), which potentially could bridge the topical content (i.e., labeled as *categories* in the framework; see Appendix A) and the ability level descriptors.

Moreover, the five level descriptors in each category in the Norwegian ICT literacy framework do not correspond with the grades in the Norwegian educational system (i.e., 13 years of compulsory education), and further explanations regarding this link are not provided in the framework (see Norwegian Directorate for Education and Training, 2012). It is stated in

the Norwegian ICT literacy framework that the ambitions specified in the framework are to be included and operationalized in the subject-specific curricula, and the task is primarily given to the subject expert groups. However, there are no explicit descriptions of *what* should be integrated into which subjects and during *which* year of schooling.

## 2.3 Assessment and Validation

Assessment is a familiar term for most people. It often connotes memories from school experiences, tests, and exams, as well as from other areas of life, such as when you try to pursue a driving license. In education, the term *assessment* reflects the wide range of methods, tools, and practices that teachers use to evaluate, gauge, and document students' learning, learning progress, academic readiness, or educational needs. Assessment has been defined in several ways. The National Council on Measurement in Education (NCME) defined assessment as “a tool or method of obtaining information from tests or other sources about the achievement or abilities of individuals.”<sup>2</sup> From this definition, it is evident that assessment is a broad term covering a range of activities in the schools. Thus, Terenzini (1989) developed an assessment taxonomy on the basis of three questions: (1) What is the purpose of the assessment (i.e., formative or summative)? (2) What is the level of assessment (i.e., who will be assessed, for instance individual students or groups)? (3) What is to be assessed (i.e., possible educational outcomes such as skills, knowledge, attitudes, and/or values)? Terenzini (1989) juxtaposed the three questions in a three-dimensional matrix, and the outcome was a categorisation of the variety of assessment approaches. Some examples of assessment approaches are placement examinations and other diagnostic tests that focus on the individual level and may intend to determine a student's readiness for learning of a particular subject or content, and/or to assign the student to the most beneficial learning sequence. Other assessments at an individual level take on the role of gatekeepers, sorting the qualified from the unqualified. Moreover, individual assessments may be aggregated to evaluate, for instance, program effectiveness or instructional quality. Along with Terenzini (1989), researchers argue that the objective of testing and assessment can often be considered in two broad categories: summative and formative assessment (Black, 1999; Andrade & Cizek, 2010). Summative assessment serves to inform an overall judgment of achievement (e.g., certificate at the end of schooling) and is designed to provide an overall picture of

---

<sup>2</sup>[http://www.ncme.org/ncme/NCME/Resource\\_Center/Glossary/NCME/Resource\\_Center/Glossary1.aspx?hkey=4bb87415-44dc-4088-9ed9-e8515326a061#back](http://www.ncme.org/ncme/NCME/Resource_Center/Glossary/NCME/Resource_Center/Glossary1.aspx?hkey=4bb87415-44dc-4088-9ed9-e8515326a061#back)

performance. On the other hand, formative assessment accounts for short-term evaluation, which may be used for the guidance of learning, mainly in daily classroom practice (Black, 1999). Keeping this distinction, the findings of the systematic review study (Paper 1) showed that the majority of the ICT literacy assessments add to the summative assessment category, as their primary function is to monitor students' ICT competences and/or provide policy makers, researchers, school authorities, schools, and/or teachers with information about the students' level of ICT literacy. Moreover, with regard to Terenzini's taxonomy, the ICT literacy assessments relate to the category *Campus and Program Evaluation* and not to the category *Gatekeeping Tests at an Individual Level* (Terenzini, 1989). The results of the assessments were most often aggregated at the group level, the focus was primarily evaluative, and the information obtained was often used by external bodies (e.g., educational authorities, educators, and policy makers), for instance for program or curriculum improvement and development.

It has been argued that assessment is closely related to teaching and learning, and that the three are inextricably linked, as each informs the others (Black, 1999; Wilson, 2005). Moreover, the importance of reliable and valid measures for supporting teaching and learning is emphasized, and assessment is seen as the bridge between instruction and learning (William, 2014). Trustworthy measures are particularly critical from a social justice perspective; principles of fairness and equity emphasize that students should be provided with equal opportunities within and across classrooms and schools (OECD, 2001). Also, from an economic rationalist view in which students' success is strongly linked to the outcomes of education, trustworthy outcomes from assessments are required (MacDonald & Brooker, 1999; Newton & Shaw, 2014). This is also accentuated from a test development perspective, in which one of the key aims is to develop measures with evidence to reflect the intended construct to a large degree. This leads to the vital role of reliability and validity investigations as a process of scrutinizing the quality of assessments.

### **2.3.1 Test Validation**

According to Newton and Shaw (2014), the terms validity and validation are often considered "two sides of the same coin," whereby validation is the process by which validity is investigated, and validity is the property to be investigated. Moreover, validity is used across different fields and contexts connoting different meanings, ranging from everyday use, to technical practices and procedures, to educational or psychological assessment.

Several assessment development frameworks (e.g., design protocols or test design principles) have been outlined to inform development and validation of reliable and valid assessments across disciplines (DeVellis, 2011; Mislevy, Steinberg, & Almond, 2003; Wilson, 2005). As shown in the left side of Figure 3, these frameworks by and large share three core principles: (1) modelling (e.g., a clear definition of what is desired to be measured); (2) operationalizing (e.g., a careful construction of the content of the scale); and (3) measuring (validation of how the scale and student scores support the intended interpretations to be made). The process of assessment development is not linear as described here, yet cyclic, emphasizing an iterative process (Wilson, 2005). Nevertheless, the linear model was used to anchor the studies to the different phases of the framework. Moreover, there is a next phase in the framework which is not included here. This phase is concerned with the outcome or results of the assessment—to inform not only the instrument development, but also the inferences made and the actions taken on the basis of the results (e.g., students' performance). Nevertheless, because the trustworthiness of the results is dependent on the quality of the test, this phase is also dependent on the initial three phases as described in the framework (Figure 3). As shown in Figure 3, Paper 1 is closely related to the first phase, *Theoretical Model*. Paper 1 is also related to the next two phases because it appraises the quality of the instruments reported in the studies included in the systematic review. The criteria for quality appraisal are concerned with how the construct is operationalized and how the measurement model fits the empirical data. Paper 2 and Paper 3 are, in addition to the first two building blocks, to a large degree concentrated around the third phase, the *Measurement Model*. Because these two papers are validation studies, the content of the tests (e.g., items or tasks; phase 2) and the underlying framework (i.e., CIL and 21st-century skills; phase 1) also play a significant role.

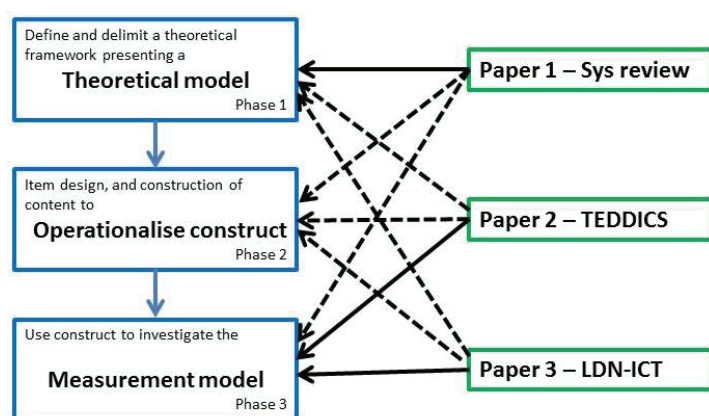


Figure 3. The assessment development framework (at the left) and the relations of the three papers to each phase of the framework

In addition to design protocols as described above, a number of well-established validity frameworks (e.g., Kane, 2006; Messick, 1995) and standards (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA & NCME], 2014; Association for Educational Assessment–Europe, 2012) exist for carrying out studies of educational assessment. Although these theoretical perspectives differ to some extent, they share common ground regarding their view on validity as a process and not a state *per se*—it is not a test that is validated. This view is in line with Messick’s (1995) definition of validity: “an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment” (p. 741). This definition indicates that all phases of the assessment development framework, for instance, are important for establishing an argument for how well the measures derived from an assessment reflect the construct of interest. This was explicitly argued by Messick (1995), as he emphasized a unified concept of construct validity that integrates the traditional single conceptions, namely content, criterion, and construct validities.

In line with this view, in Paper 1, a number of indicators (which refer to the evidence for the reliability of the test instruments and how a validity argument was created; see Appendix B in the systematic review) were identified. These indicators were based on acknowledged literature (AERA, APA, & NCME, 2014; Newton, & Shaw, 2014; Zumbo & Chan, 2014), represent an integrative view on validation consisting of generic qualities, and are not specific to certain methodologies or research traditions. The main objective of the indicators was to facilitate the examination of the reported quality of the assessments, including aspects such as development procedures, psychometric properties, and explicit information about the content and intended use of results. Thus, the indicators align well with the phases of the assessment development framework. Some indicators were specifically included because assessments of ICT literacy differ from many other assessments. For instance, most of the tests are computer-based and include a variety of tasks/items presented to the test-takers. For evaluation of the authenticity and the interactivity of the task/item design, one indicator was developed which was used to inspect the task/item design according to how the information (e.g., dynamic, static) was presented, what the test-taker had to do to find the right answer, and how the response was given (e.g., multiple-choice, performance).

Furthermore, the indicators developed in Paper 1 apply to Paper 2 and Paper 3, because these are validation studies aiming at investigating two different constructs.





## 3 Research Methods and Methodological Considerations

In the three studies, different research methods were applied to address the research objectives posed in the dissertation. The primary focus was not on the research methodology; rather, the objectives of the studies led to the choice of method. In Paper 1, a systematic review methodology was followed. In Paper 2, exploratory structural equation modeling (ESEM) was used, and in Paper 3, item response theory (i.e., unidimensional and multidimensional Rasch modeling) was used. The methods used in Papers 2 and Paper 3 are to a large degree described in the papers, yet the reasoning behind *why* these methods were chosen could not be provided sufficiently due to journal restrictions, and will therefore be briefly described in this chapter. Moreover, the methodological considerations and elaborations related to the systematic review methodology (Paper 1) will be outlined in the following because these could not be detailed due to restrictions of the format of systematic review papers. Furthermore, Study 3 is concerned with the LDN-ICT test, which was translated, adapted, and revised. The paper investigates the reliability and validity of the test, but the nature of the article format did not allow for a presentation of the preliminary process of preparing the test for Norwegian students. In order to document this crucial and laborious part of the work, some more details are provided.

### 3.1 Systematic Review (Paper 1)

This section provides a rationale for conducting the systematic review, followed by a description of the main differences between a traditional and a systematic review of literature. Finally, the systematic review processes are elaborated.

#### 3.1.1 Rationale for Conducting a Systematic Review

I started my doctoral studies with the aim of assessing Norwegian lower secondary students' ICT literacy. This choice of theme (research area) led to several questions: What is ICT literacy? Why should one care to measure it? How can it be measured? And how is it *being* measured? The latter became the driving force for considering a systematic review. Petticrew and Roberts (2006) provided five arguments for when one should conduct a

systematic review. Three of them were particularly relevant for Paper 1: (1) when it is known that there is a wide range of research on a subject but where key questions remain unanswered, (2) when a general overall picture of the evidence in a topic area is needed to direct future research efforts, or (3) when an accurate picture of past research and past methodological research is required to promote the development of new methodologies (p. 21). In addition to these mere general views, several reflections with regard to ICT literacy assessment came into play, such as the comprehensiveness of the ICT literacy term, which is broad and consists of several competence areas and competences (Calvani, Cartelli, Fini, & Ranieri, 2008; Eshet-Alkalai, & Amichai-Hamburger, 2004; Ferrari, 2013) and which may not be easily measured in one test (Calvani et al., 2008). Also, there is a risk that by using different terminology, the existing literature may not be fully connected. Moreover, the assessment of ICT literacy was to some extent blurred by the large diversity of assessment instruments. Many of these are based on self-reports and self-efficacy measures, implying that students' own perceptions and judgments are used as proxies for their actual ICT literacy (Aesaert et al., 2014). Studies comparing students' self-reported competence with test scores have shown that self-reports do not portray students' true competence (Law et al., 2009; Vonkova & Hrabak, 2015). Another source of diversity between existing ICT literacy assessments is the degree to which the tests capture actually performed competence, in other words the degree to which the tests are based on authentic tasks conducted in test administration software representing the dynamic and interactive nature of regular software. An initial literature search revealed limited access to handbooks and systematic reviews considering the research area described above, yet a number of primary research studies exist. Altogether, a systematic review seemed important for describing the state of the art of this field, revealing research gaps, and outlining needs for future research.

### **3.1.2 Traditional Versus Systematic Reviews**

Gough, Oliver, and Thomas (2012) emphasized that “any individual research study may be fallible, either by chance or because of how it was designed and conducted or reported” (p. 3). Furthermore, they argued that traditional literature reviews present findings related to a topic of interest and tend to summarize what is known about a topic in an unbalanced way by including only the selection of literature most closely related to the researcher's own conceptualization and theoretical framework, leading to the intended or unintended exclusion of literature from slightly different fields or research approaches.

Moreover, traditional reviews do not explain criteria used to locate and include studies, which makes it difficult to prove the appropriateness of the decisions and whether they were applied thoroughly and consistently. The seminal paper by Kahneman, Slovic, and Tversky (1982) argued that people often select, evaluate, and remember information that supports their individual preferences. They accentuated that we fail to look for evidence that disconfirms our pet hypotheses, and we do not spot errors in our own reasoning. Moreover, we examine more critically the evidence that contradicts our own views, and vice versa, it is easier to accept evidence in line with our own views (Tversky & Kahneman, 1974). Therefore, the systematic review method stresses the need for using explicit, rigorous, and accountable methods (Gough et al., 2012). Also, involving several persons (and stakeholders) in the review process has been emphasized as vital for the quality of the review, and researchers suggest that all studies included in a review should be independently reviewed by at least two researchers (Gough et al., 2012).

Systematic reviews have their origin from clinical medicine, where they have a long tradition (Jesson et al., 2011). Other research areas, such as education, also have started to adopt systematic reviews as a way to summarize the available knowledge about specific issues or research questions. The Cochrane Collaboration defines a systematic review as follows: “A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions madet” (Green et al., 2015). Systematic reviews vary to a large degree depending on the research questions addressed (Gough et al., 2012; Petticrew & Roberts, 2006). However, they share common principles, which are described in the following section.

### **3.1.3 Common Steps in Systematic Review Methodology**

In Paper 1, several processes were followed as suggested in many systematic review methodology guidelines (Gough et al., 2012; Petticrew & Roberts, 2006). As shown in figure 4, there are six phases which are common to most systematic reviews. The *first phase* is concerned with defining the research problem and designing the review by developing a protocol. This part is essential to the systematic review, as it enables the researcher to conduct a rigorous and transparent review by pre-defining and clarifying questions, for instance the conceptual framework, eligibility criteria, and search strategy. For the systematic review, this

was a critical phase because the field, as stated, is populated with literature using different terminology which refers to slightly different theoretical traditions. The protocol may portray similar function as a well-planned study design in a primary research study, reflecting the theoretical considerations taken into account to shed light on the research question the study has addressed. The protocol of Paper 1 can be found in Appendix B. The *second phase* is related to conducting the actual search by applying the search strategy to the pre-identified databases and other sources. This phase also includes the screening of the search output in line with the pre-defined inclusion and exclusion criteria.



Figure 4. Six phases typically included in a systematic review

The *third phase* is concerned with descriptions of study characteristics, and the aim of this phase is to extract the information needed to address the research question(s) of the review. In Paper 1, the information was extracted following a coding scheme which was developed in an iterative process. On the basis of the initial coding, the coding scheme was further developed and revised and used across all included studies to extract the relevant information. *Phase four* is concerned with quality appraisal. For this phase, a considerable amount of theoretical work was done to generate the criteria for appraising the quality of the studies included in the review. The *fifth phase* is labeled synthesis, and it is concerned with using the extracted information together with the theoretical framework, coding schemes, and quality indicators to address the research questions posed. In this phase, the results are discussed and conclusions are drawn. The *last phase* is related to dissemination of the findings, which was achieved through presentations at conferences and later as the initial manuscript went through a peer-review process before being published in an international journal. The process of systematic reviews, as described here and indicated in Figure 4, may give the impression that the methodology is entirely straightforward, which is not the case. The process is much more complex and iterative; therefore, arrows could have been added from each phase and back to the previous phase or even to the phases before. That may have shown the complexity involved in conducting a systematic review, but it may have made it difficult to grasp the core features of a systematic review methodology.

## 3.2 TEDDICS (Paper 2)

The main aim of Paper 2 was to validate the TEDDICS scale. A number of approaches and methods can be utilized for investigating the reliability and validity of a construct. For the TEDDICS scale, which was developed within the ICILS, initial piloting of the items and the scale was conducted as part of the field trial. On the basis of the main data collection, the TEDDICS scale was reported as a single composite measure (Fraillon et al., 2014). Paper 2 re-investigated the full set of Norwegian ICILS data with respect to the theoretically expected factorial structure, generalizability, and relations to other constructs by using exploratory structural equation modeling (ESEM).

### 3.2.1 Exploratory Structural Equation Modeling (ESEM)

The ESEM methodology was used due to its applicability in studies in which the theoretical assumptions on the multidimensionality of the construct are fairly strong, where invariance of the scale across groups needs to be ensured, and when relations to other constructs are to be investigated. In contrast to confirmatory factor analysis (CFA), ESEM allows a more flexible representation of the factor structure, which can provide a more accurate representation of reality (Marsh, Morin, Parker, & Kaur, 2014).

ESEM was developed by Marsh et al. (2009) and is described as “an overarching integration of the best aspects” of CFA, structural equation modeling (SEM), and exploratory factor analysis (EFA) (Marsh et al., 2014, p. 85). There are several advantages of using ESEM. First, even though CFA is a commonly used method and has largely superseded EFA, it has restrictions when it comes to multidimensional constructs. CFA relies on the assumption that each item loads only on one factor (Marsh et al., 2014). Hence, this restriction may challenge the situations in which well-defined, *a priori* assumptions about the factor structure exist. This was the case with the TEDDICS scale, which was aligned with the CIL framework (see section 2.2.2 for details about the CIL framework). Moreover, Marsh and colleagues (2014) have argued that CFA of multidimensional, mainly self-reported constructs often fails to meet standards of measurement such as an acceptable goodness of fit (Marsh, Hau, & Grayson, 2005), measurement invariance (Millsap, 2011), lack of differential item functioning (Camilli, 2006), and discriminant validity (Campbell & Fiske, 1959). Second, EFA in which all cross-loadings are freely estimated restricts the test of measurement invariance (Marsh et al., 2014), which is critical to make comparisons across groups of

persons such as gender and age groups (Millsap, 2011). Additionally, the researchers argued that it is challenging to incorporate latent variables that follow a structure with cross-loadings (derived from EFA) into subsequent analyses such as testing relations with other constructs. Moreover, given the flexibility of EFA, latent variables corrected for measurement error cannot be incorporated, and therefore constructs identified through EFA have to be converted to suboptimal scale or factor scores (Marsh et al., 2014).

Given the limitations of CFA and EFA, as described by Marsh et al (2014), and the *a priori* knowledge about the dimensionality in the frameworks that are described in the theoretical background section, ESEM was selected as competing with the CFA approach in Paper 2. The model fit comparisons were conducted to address the first research question regarding the anticipated dimensionality of the model. Moreover, the ESEM model showed better fit and confirmed the assumptions about the three distinct but related factors. The ESEM model was further used for examining the measurement invariance of the scale across gender and main subject groups.

### **3.3 LDN-ICT (Paper 3)**

The LDN-ICT test was developed within the ATC21S project, and the development of the initial test was led by the BEAR center at UC Berkeley. The initial version of the test was in English, and a small pilot had been conducted at the time the collaboration with the BEAR center was initiated (the pilot has been published in Wilson & Scalise, 2015). One option was to keep the English version of the test to assess the Norwegian students' ICT literacy in digital networks. Nevertheless, the chances for getting biased results were high given that the test probably would favor students with higher proficiency in English, or stated differently, a test in English would have introduced a major source of construct irrelevant variance. Besides, the Norwegian students probably would not have understood some of the task instructions, due to language and cultural differences. Therefore, the LDN-ICT test was translated, adapted, and revised to fit the Norwegian language, student, and school culture. In the sections below, the process of translation, adaptation, and revision is briefly described because this could not be sufficiently outlined in the paper. Moreover, the background questionnaire which was developed to accompany the test and gather more information about students' beliefs and characteristics is presented and discussed.



### 3.3.1 Adapting the LDN-ICT Test to the Norwegian Context

The translation of the test was done in an iterative process. The first translation was done by me and a student assistant. Colleagues from the department helped by taking the test to provide feedback on their understanding of the tasks, and further changes were made. One colleague who works on developing the national test in reading helped with reaching consistent and correct use of language in the LDN-ICT test. Think-aloud protocols and interviews with eight students taking the test were used to investigate and improve the functionality, user experience, and content of the translated test.

The LDN-ICT test contains three scenarios (i.e., modules), each of which includes a number of tasks and relates to three different contexts (Scenario 1: *Arctic Trek*, Mathematics and Natural Science; Scenario 2: *Human Legacy*, Social Science and Arts; and Scenario 3: *Second Language Chat*, Language). All three scenarios were translated into Norwegian, and further revisions were implemented. In addition to translation, relevant Norwegian web sources that were compatible with the web pages in English had to be identified; where this was not possible, equivalent resources had to be created.

The aim was to keep the translated version as close as possible to the original test. Nevertheless, the translation of the *Human Legacy* scenario required larger changes and replacements of some tasks. This scenario was framed as part of a poetry work unit, in which students were supposed to read and analyse well-known poems (Wilson & Scalise, 2015). The English poems were replaced by Norwegian poems, which were familiar to Norwegian students and generally used in classroom settings. Furthermore, authentic activities that typically occur in classrooms were incorporated into the tasks; for example, the students were asked to express their interpretation of the moods and meanings of the poem, and whether the YouTube video they watched was an adequate interpretation of the poem. Due to issues with the software licence of the software *Webspiration* outside the United States, the *Human Legacy* scenario was supplemented by a new task, with the aim of investigating students' collaborative problem-solving competences. The task required students to meet their group by accessing a web page link through the test to the online software *CoSketch*. This software consists of a drawing tool and an embedded chat. Students were asked to sketch a drawing together in groups of three students. The resultant drawing was to express their interpretation of the poem they had just read and watched a video about. Because *CoSketch* is shared software, it allows for only one drawing at a time. Thus, the students had to draw together and collaborate through the chat in the same software. At the end, the students were supposed to

save and upload the collaboratively sketched drawing along with the group conversation (the chat).

Think-aloud protocols of an additional 12 students were conducted at a later stage to evaluate the collaborative problem-solving task, and investigate to what extent the task facilitates and captures students' abilities to interact with each other and solve a problem within an assessment. Moreover, the aim was to investigate to what extent the processes described in the underlying framework occurs, as well as to use the information to develop a scoring rubric. The video data of the students' think-aloud protocols, while they were taking the test, have been coded, and initial analysis has been conducted. The preliminary results show that the task is largely successful in capturing student-student interactions and collaborative problem solving in a digital environment. However, due to the time restrictions, this study was not finalized within the completion of this dissertation, and the data collected will be subjected to analysis as part of a follow-up project.

### **3.3.2 Questionnaire Development**

A questionnaire was developed to gather more information about the students and to further investigate the validity (i.e., external validity) of the LDN-ICT test. After taking the test, the students were directed to the online background questionnaire, which consists of constructs related to students' use of ICT, ICT self-efficacy, collective efficacy, perceived usefulness, academic aspirations, learning strategies, mastery orientation, and background variables such as gender, ethnicity, SES, and self-reported grades in different subjects. Each sub-construct was included on the basis of literature review in the field of ICT literacy assessment in general and research on assessment of group work. A varying number of items were developed and assigned to each sub-construct, and the items reflected to a large degree the content of the test. Some of the sub-constructs were analyzed and used in Paper 3, whereas others have not been investigated yet. Note that not all sub-constructs were developed from scratch. For instance, the sub-constructs on ICT use were revised from a previous questionnaire (developed by me and another colleague) and updated to include items which refer to smartphones and tablets and in particular address communication and collaboration to better align with the construct measured in the test.



### 3.3.3 Item Response Theory

Several steps were taken to ensure the validity and reliability of the test. First, after several rounds of translation and revisions, students' think-aloud protocols while they were taking the test were used to ensure that the students understood the tasks as they were intended and that the software functioned properly. Furthermore, focus-group interviews were administered to collect information about students' perceptions and experiences with sitting for the test. After the pilot, the test data were used to statistically investigate the reliability and validity of the test, which is reported in Paper 3. A number of statistical approaches can be taken to address validity and reliability of assessments. Classical test theory (Christensen & Knezek, 2008) is still commonly used for this purpose, but several shortcomings with this approach have been identified, for instance that the analyses are performed on the test as a whole rather than on single items (Hambleton & Jones, 1993), and even though it is possible to generate item statistics such as item-to-total correlations, these still only apply to *that* particular group of respondents and to *that* particular collection of items (de Ayala, 2013; Embretson & Hershberger, 1999). Conversely, in IRT, the estimation of respondents' abilities is independent of the actual items included in the test and vice versa (Wilson, 2005). Another advantage of IRT is that it provides better estimates of measurement accuracy by acknowledging that the test gives more information about respondents with abilities that are well targeted by the items (Hambleton & Jones, 1993; Thomas, 2011). Specifically, item fit (i.e., measure of the discrepancy between the observed and theoretical scores; Wu & Adams, 2013), local independence (the responses to the items are independent; Smith, 2002), and differential item functioning (DIF) (i.e., measurement bias for an item that occurs when different subgroups (e.g., gender) in a sample have different probability of giving a correct response given equal levels of proficiency; Tennant & Pallant, 2007) are useful approaches to investigate the validity of a scale. Furthermore, it has been accentuated that these benefits are particularly important with regard to the transferability and reuse of assessment instruments, and they render the possibility to design instruments with specific characteristics (de Ayala, 2013). Intrinsically, given that the LDN-ICT test was translated, revised, and further developed, the main goal was to investigate its feasibility as an instrument for measuring students' ICT competences in digital networks. For this purpose, the IRT approach and in particular Rasch modelling was chosen as a promising method (Aesaert et al., 2014). Furthermore, given the *a priori* knowledge about the dimensionality in the frameworks, the multidimensional Rasch model was used to investigate the data, as it is considered the

simplest and most parsimonious model within the family of multidimensional IRT models (Adams, Wilson, & Wang, 1997).

### **3.4 Reflections on Research Credibility**

Research credibility is concerned with the close relation between reliability and validity (Kleven, 2008; Shadish, Cook, & Campbell, 2002). Reliability refers to the degree to which the findings of a study under the same conditions produce the same results (Abercrombie, Hill, & Turner, 1984; Silverman, 2006). Validity refers to “an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment” (Messick, 1995, p. 741). In assessment, reliability is often seen as an integral part of validity (Wilson, 2005). Moreover, researchers have argued that reliability (consistency) is a necessary condition for ensuring high measurement quality—but not sufficient (Newton & Shaw, 2014). The main argument for this view is embedded in the fact that an assessment instrument may measure something wrong—but consistently (e.g., two watches can be consistent but be six hours slow, Newton & Shaw, 2014, p. 14). Hence, they concluded that a test must be reliable to be valid, but the opposite does not apply, and therefore reliability is considered one facet of validity.

Moreover, both validity and reliability are generic meta-concepts, and in discussing the validity and reliability of the results and claims made from a specific study, the concept needs to be contextualized to take into account the research design, the methods used, and the overall purpose of the study. Therefore, because different methodological approaches were taken in the three studies, slightly different criteria apply. Hence, an overall discussion of the validity and reliability for the thesis as a whole is not feasible or useful. Instead, it seems more relevant to quest the trustworthiness of each of the studies separately. Several steps were taken to establish the rigor of the work presented in the dissertation. One of them was transparency, which means that the researcher aims at being explicit about what methods and analytical tools were used to collect and analyze the data, thus making it possible for others to replicate the work (Lincoln & Guba, 1985).

Ercikan and Roth (2006) argued that the research questions posed should dictate and determine the most appropriate modes of inquiry, and they discussed that polarizing educational research into quantitative and qualitative or subjective and objective is neither

meaningful nor productive. Therefore, they proposed an integrative framework in which research studies are placed along a continuum of low-level inference to high-level inference, and they claimed that the traditional distinctions between qualitative and quantitative approaches are instead located at different sides of the same scale. The low-inference side is characterized by contingency, particularity, being affected, and concretization. The opposite side is characterized by standardization, universality, distance, and abstraction (Figure 5). An evaluation of these categories justifies placing the first two studies (Systematic review and TEDDICS) at the right side of the scale, indicating findings/knowledge at the high-inference level. In establishing a validity argument for the claims made in these studies, it was of vital importance to document how well the results could be generalized from the sample to the population.

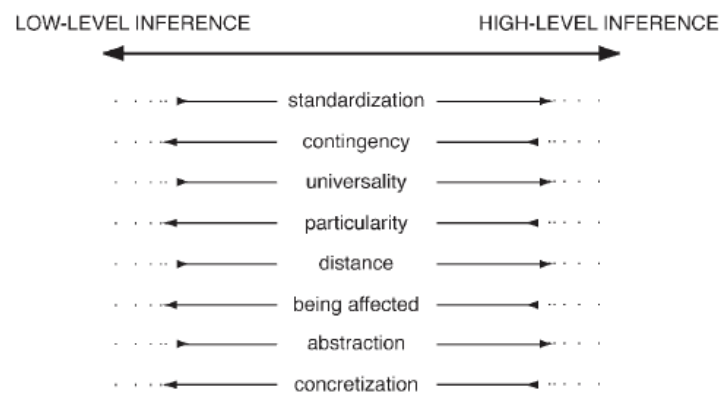


Figure 5. Continuum of low-level-inference to high-level-inference research and associated tendencies for knowledge characteristics along eight dimensions (from Ercikan & Roth, 2006)

Paper 3, on the other hand, is closer to the middle of the scale because it had a small sample size and the students were purposefully sampled. Besides, in the preparation of the LDN-ICT test, think-aloud protocols and interviews were conducted, which would have been placed more at the left side of the scale (low-level inference), yet these were critical for ensuring that the students understood the tasks in the way they were intended, and making sure that other issues did not come into play when the students took the test. In establishing a validity argument for this study, it was vital to ensure that the revised LDN-ICT test measures the intended construct, which was done with different approaches (i.e., translation and revisions of the original test by involving different people, think-aloud protocols, interviews, and a pilot). Because the focus in this paper was on the validation of a relatively new test,

containing features (e.g., facilitates synchronous collaboration between students, open access to the Internet) that had not been largely integrated in ICT literacy assessment earlier (as shown in Paper 1), the focus was on the test and not on the sample. Thus, the inferences drawn on the basis of the study are related to further refinements of the test.

Shadish et al. (2002) emphasized that in addition to a well-defined research design and process, the validity of the inferences is important to enhance the credibility of the research. With regard to this, the results and inferences of the three studies in the dissertation were discussed against existing theory and with researchers in the field (e.g., at conferences). In addition, the peer-review process in academia contributed to this (e.g., conference presentation proposals and journal submissions).

## 4 Results

The detailed findings are presented in each of the three papers. However, a brief summary of the main findings will be provided in the next section, before a discussion of the results and implications for the field in the subsequent chapters.

### 4.1 Summary of Main Findings

#### *Systematic Review (Paper 1)*

Paper 1 investigated the first research objective: What is the current state of the art of the field of assessment of primary and secondary students' ICT literacy?

The systematic review identified 38 tests developed in 17 different countries. A limited number of the tests include cross-country comparisons or measures of trend. The studies were published between 2000 and 2014, and most of them were conducted with lower secondary school students (grades 7–10) and were associated with a framework of ICT literacy. Most of the tests comprised multiple-choice-item designs and were evaluated by quantitative methodology. To identify which facets of ICT literacy the assessment instruments aimed to measure, the reporting of the content of each test was scrutinized against the revised DIGCOMP framework. Most of the tests measured the competence area *Information*. The competence areas *Content Creation*, *Communication*, and *Technical Operational Skills* also were well covered, whereas a few tests measured the areas *Safety* and *Problem Solving*. Detailed analysis of each competence area showed that the competences measured within the area *Communication* were concerned with students' knowledge *about* digital communication and collaboration and not *how* they communicate and collaborate digitally. In particular, a lack of tests that measured *Synchronous Digital Communication and Collaboration*, *Safety*, *Programming*, or *Problem Solving* was identified. Furthermore, the reporting of the quality of the tests and test development process was investigated against 15 indicators, supporting the reliability evidence and the creation of a validity argument. The majority of the tests reported fewer than ten indicators, suggesting that the reported information about the quality of the instruments and their intended interpretations was by and large incomplete across the reviewed papers. This finding is critical and questions the trustworthiness of the tests and the credibility of the results acquired with a specific test. The systematic review provides an

overview of the field, and by identifying the research gaps, it attempts to outline possible future directions.

### ***TEDDICS (Paper 2)***

Paper 2 investigates the second research objective: To what extent can teachers' emphasis on developing students' digital information and communication Skills (TEDDICS) be measured with high quality?

The results indicate that the hypothesized three-factor structure of TEDDICS was supported by the ESEM and showed improved model fit compared to a CFA model. The results suggested an overlap among the three factors, as manifested in significant cross-loadings. The investigations of the external validity of the construct indicated positive and significant relations between the TEDDICS factors and ICT self-efficacy, ICT use, and perceived usefulness of ICT. These results indicate that, for instance, teachers who believe in their own competences of using ICT (self-efficacy) or view ICT as important (perceived usefulness) also emphasize developing their students' ICT literacy more. Moreover, evidence for the highest level of invariance (strict invariance) was ensured, and consequently comparisons of the factor means between male and female teachers and across the different main subjects could be employed. Further analysis showed significant mean differences in TEDDICS in favor of teachers in the humanities, language, and art, yet no significant gender differences were found.

### ***LDN-ICT (Paper 3)***

Paper 3 addresses the third research objective: To what extent can the validity of the Learning in Digital Networks-ICT literacy (LDN-ICT) test be provided?

The LDN-ICT test is a performance-based online test with open access to the Internet, which measures students' competences in handling digital information, content creation, synchronous digital communication, and collaborative problem solving. These competences are operationalized in four strands (i.e., dimensions). The results showed that the hypothesized four-dimensional Rasch model fit the data better than a unidimensional Rasch model, which indicates that the underlying framework could be operationalized. DIF analysis was applied to investigate whether items function differently across groups (i.e., students' gender and SES). Five items which indicated DIF across students' gender and SES were removed, and evidence for an invariant construct was ensured. Thus, further comparisons

across these groups could be made, and positive correlations between SES and the four dimensions were revealed, yet no gender differences were indicated. The correlations between the LDN-ICT test and collective efficacy and perceived usefulness of ICT were positive, nevertheless insignificant. Positive and statistically significant correlations between the four dimensions of LDN-ICT and ICT self-efficacy and academic aspirations were revealed, indicating evidence for the tests' external validity.





## 5 Discussion

Each of the three papers in the dissertation discussed the results and addressed the implications for the field following the narrow focus in each paper. In this section, the results are discussed in the light of the overarching aim of the dissertation, which is to investigate the assessment of ICT literacy in the educational system by examining the positions and perspectives of different actors (teachers and students) and practices (assessment instruments) to portray how educational systems can monitor and support the development of students' ICT literacy through relevant, purposive, and high quality assessments.

The introduction chapter aimed at shedding light on the importance of ICT literacy in education, particularly for students (e.g., motivation and equal opportunities to learn in the 21st century). The responsibility of teachers as facilitators of students' learning of ICT literacy, as well as the critical role of ICT literacy assessments to monitor and eventually seek to realize the objectives of this study was described. The theoretical and methodological reflections in Chapters 2 and 3 clarified and further explained some of the choices made in the dissertation and facilitated the discussion. This chapter provides an overview of the dissertation by revisiting the curriculum model, and it addresses an alignment among the three papers and the overarching aim. Furthermore, some of the findings presented in Chapter 4 are discussed across the three papers; when appropriate, implications for policy and practice are suggested.

### 5.1 An Overview of the Dissertation

An integrated view on ICT literacy is emphasized in the dissertation. The curriculum model is used to frame the three papers and to emphasize that the works presented in the dissertation are conceptually related, and as a whole they shed light on an integrated view of ICT literacy. The curriculum model has been emphasized in studies of ICT literacy to take into account the complexities involved when studying ICT literacy at a system level (Markauskaite, 2006; Voogt & Roblin, 2012). The model consists of three levels: the *intended*, the *implemented*, and the *attained* curriculum, and the alignment between the research objectives and the model was presented in Chapter 1.3. The *intended* level in the curriculum model is concerned with what national and international educational policies identify and value as important ICT literacy competences (e.g., policy documents, explicit and

formalized curriculum), and that was mostly addressed in Paper 1 (Figure 6). The *implemented* level refers to the educational practices at the school and/or classroom level, and it is usually described in terms of the efforts extended to help students attain the curriculum goals (Mullis et al., 2009). As shown in Figure 6, this level was investigated in Paper 2. The *implemented* level is dependent on the curriculum and/or educational standards (i.e., *intended* level), as they direct the arrangements of teaching and learning. The *attained* level refers to what the students have learned as a result of their educational experiences and their attitudes toward the subjects (Markauskaite, 2006), and Paper 3 provides one tool for inspecting this level. As shown in Figure 6, the *attained* level is also affected by the *intended* and *implemented* levels, because the *intended* curriculum is relevant for students' learning goals and what should be assessed. Moreover, what students learn at school (i.e., the *implemented* level) would most likely have an effect on their learning outcomes (e.g., achievement on the ICT literacy test). In reciprocity, the outcome of students' achievement may affect the *intended* and the *implemented* levels, because empirical investigations have been emphasized as critical to develop higher-quality curricula (Thijs & van den Akker, 2009), and the strengths and weaknesses of students' competences may influence teachers' instructional practices. Hence, the three levels interact (Goodlad et al., 1979). This is also illustrated in Figure 6, which provides an overview of the dissertation. On the basis of these interrelations, two general themes were identified across the three studies: ICT literacy frameworks and assessment, which will be discussed in the subsequent sections.

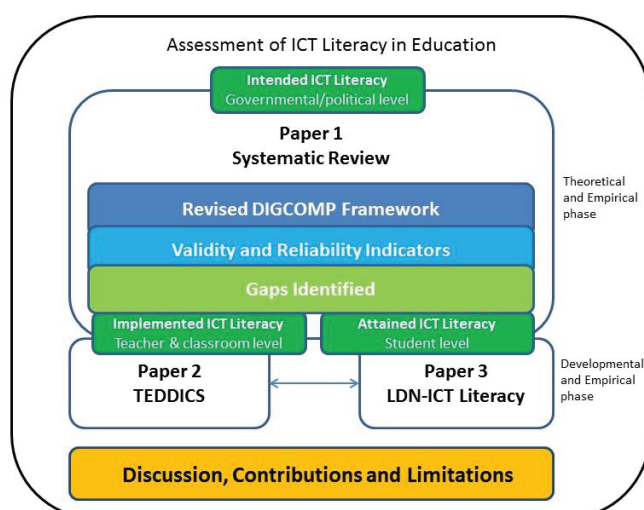


Figure 6. An overview of the dissertation, including the interrelation between the three papers and the curriculum model, as well as how they address the overarching aim of the dissertation

## 5.2 ICT Literacy Frameworks

The frameworks of ICT literacy have been emphasized in this work because they are concerned with what national and international educational policies identify and value as important ICT competences. Besides, as described in the assessment development framework (section 2.3), the theoretical frameworks play an important role because these are used to define and delimit test content and prior expectations about students' or teachers' conceptual understanding, which eventually guides item development. Therefore, taking into account the frameworks underlying the three studies is important. In addition, the Norwegian curricular context is emphasized because the respondents (i.e., teachers and students) in Paper 2 and Paper 3 were part of the Norwegian educational system.

The policy documents or explicit and formalized curriculum documents usually also include normative arguments for why these aims or intentions have been selected. However, these normative arguments are not extensively presented and discussed in the dissertation.

In the subsequent parts, the most important findings related to the frameworks are discussed under three sub-categories: alignment of the theoretical frameworks, operationalization of the ICT literacy frameworks, and the Norwegian ICT literacy curriculum.

### 5.2.1 Alignment of the Theoretical Frameworks

The three international frameworks (i.e., DIGCOMP, CIL, and 21st-century skills) underlying the three studies were, as shown in Chapter 2, developed within different, larger international enterprises, and they serve slightly different aims. However, all three are concerned with ICT literacy and address assessment as a vital aspect. While the CIL framework is rather narrow compared to the revised DIGCOMP framework, the 21st-century skills framework is fairly broad. Nevertheless, both CIL and the 21st-century skills framework provide detailed descriptions of the content. Both identify and delineate what is meant by ICT literacy, as well as group and define the underlying competence areas and the related competences. Taking into account the different purposes, common facets of the frameworks could be identified. For instance, using the revised DIGCOMP as a baseline, the conceptual content (i.e., competence areas and competences) in the other two could be identified, allowing for comparisons among the three. These findings are in line with previous research which revealed that the ICT literacy frameworks converge on a common set of 21st-

century competences, although they use different terminologies, grouping, and categorization procedures (Voogt & Roblin, 2012).

Moreover, the revised DIGCOMP framework was used in the systematic review to appraise, evaluate, and distinguish which facets of the framework the different tests measure. The results showed that the revised DIGCOMP framework has the depth and breadth to cover all 34 tests (and their underlying frameworks) that were included in the review. These findings further consolidate the identification of common features in the theoretical frameworks (Gallardo-Echenique, de Oliveira, Marqués-Molias, & Esteve-Mon, 2015), and it became evident that the revised DIGCOMP can serve as a reference model. In conclusion, the systematic review contributes to the literature by showing that the different concepts, labels, and frameworks are comparable (Ala-Mutka, 2011; Calvani et al., 2012; Hatlevik & Christophersen, 2013). This may lessen some of the confusion caused by the use of different labels when addressing ICT in education.

Due to the compatibility of the frameworks shown, further discussions and implications in the dissertation can be elaborated without reference to the specific frameworks. In the following, reference is accordingly given to the concepts as they are organized and labeled in the revised DIGCOMP framework.

### **5.2.2 Operationalization of ICT Literacy Frameworks**

The revised DIGCOMP framework was used as a reference point to investigate and document which competence areas and competences are operationalized in the tests included in the systematic review. The systematic review revealed that some competences are more frequently included than others. For instance, the competence area *Information* stands out as an area covered by most tests in the field. The competence areas *Content Creation*, *Communication*, and *Technical Operational Skills* are included to some extent, whereas few tests cover the competence areas *Safety* and *Problem Solving* (see Table 1 for an overview of the competence areas and competences). Moreover, a very limited number of tests operationalize competences such as *Synchronous Digital Communication and Collaboration*, *Protecting Devices*, *Health and Personal Data*, *Programming*, and *Problem Solving*. Hence, the results exposed an imbalance between the theoretical frameworks and the operationalization of these and reveal that assessment is lagging behind (Quellmalz, 2009).

The imbalance between the ICT literacy frameworks and the operationalization of these in the tests revealed in Paper 1 was investigated on a theoretical basis. This finding is

followed up in Paper 2 and Paper 3, which in this context can be seen as cases of operationalization of two ICT literacy frameworks and which empirically investigate the alignment with the underlying frameworks.

### ***Two Examples of Operationalization of ICT Literacy Frameworks***

Previous research on assessment of ICT literacy has to a large degree reported ICT literacy as a composite scale (Kim & Lee, 2013; Kim, Kil, & Shin, 2014; Zelman, Avdeeva, Shmis, Vasiliev, & Froumin, 2011). Yet these studies referred to multidimensional constructs. Some studies that empirically investigated the underlying multidimensional construct could not prove the dimensionality in the theoretical frameworks (Aesaert et al., 2014; Claro et al., 2012; Senkbeil, Ihme, & Wittwer, 2013). With regard to this, an adjustment to the underlying multidimensional construct was not obvious because findings from previous studies proved that the dimensionality in the theoretical frameworks can be difficult to operationalize. However, there are studies that were successful with providing evidence for the adjustment to the underlying multidimensional construct, for instance studies that measured teachers' computer self-efficacy (Scherer & Siddiq, 2015) and students' ICT literacy (Huggins, Ritzhaupt, & Dawson, 2014). In line with this, Paper 2 and Paper 3 empirically investigated the dimensionality in the underlying frameworks.

The TEDDICS construct is concerned with the extent to which teachers emphasize developing their students' digital information and communications skills. In this measure, the competence area *Information* is covered, as well as some competences related to the area *Communication* (with reference to the revised DIGCIMP framework). This scale was chosen as an object of investigation primarily because it investigates the teacher perspective and secondarily because the scale was to a large degree aligned with the content of the tests that measure students' ICT literacy (in particular the competence area *Information*). This is not surprising, because it was developed within the large-scale international assessment ICILS, which mainly focuses on students' CIL as a subset of competences (Fraillon et al., 2014). In the ICILS report, the TEDDICS scale was reported by one overall score. Nevertheless, the empirical investigations of the TEDDICS scale (Paper 2) revealed that the underlying framework could be identified, and a multidimensional model consisting of the three factors fit better than a unidimensional model.

The findings of the TEDDICS study revealed that teachers to a great extent emphasize developing students' competence in searching for digital information. In contrast, a fairly low

level of emphasis on evaluation of digital information was indicated. These results are in line with research on students' ICT competences which disclosed that students struggle with evaluation of digital information (Claro et al., 2012; Kuiper et al., 2005; Strømsø & Bråten, 2014). These results emphasize the importance of taking the factor structure of the TEDDICS instrument into account. This enables the reporting of the separate dimensions, providing more detailed information that can be valuable for improved understanding of teaching and the alignment of the *implemented* and the *intended* curriculum.

The third paper (LDN-ICT) investigated aspects of ICT literacy that have barely been operationalized in tests, although they are identified as critical for students to function in work life and the knowledge society (Binkley et al., 2012). The LDN-ICT test operationalized competences within the competence areas *Information* and *Content Creation*, which are well covered as suggested by the systematic review findings. In addition, the LDN-ICT test measures competences within the competence areas *Communication* and *Problem Solving*, which have been scarcely operationalized in previous studies (Quellmalz, 2009). The empirical investigations of the LDN-ICT test revealed that a Rasch model with four dimensions had a significantly improved model fit as compared to a unidimensional model, lending support to the theoretical framework.

Moreover, the analysis showed that the students scored higher on the competence areas *Information* and *Content Creation* (labeled as *Consumer in Networks* and *Producer in Networks* in the LDN-ICT paper) compared to the competence areas *Communication* and *Problem Solving* (labeled as *Social Capital* and *Intellectual Capital*). These detailed results (based on the evidence for the four-dimensional Rasch model) indicate that the students need more training in some competence areas compared to others; this could be used to further inform instructional practice. Therefore, an alignment with the underlying framework and empirical investigations of the multidimensionality of it seem to be a fruitful approach. A reporting of the overall score on the ICT literacy tests provides essential information and could be useful to compare groups of students (e.g., age, gender, SES, ethnicity). However, a detailed reporting at sub-construct level provides detailed information which could be used for designing classroom activities to increase students' learning of these competences, as well as by teachers for formative assessment purposes.

In sum, even though further refinements of the TEDDICS and LDN-ICT scales are recommended (e.g., need for more items in some sub-dimensions), the studies provide



statistical evidence for the multidimensionality of the scales aligned with the underlying frameworks. And the advantages of a multidimensional view were demonstrated above.

As a conclusion, both from a theoretical (Paper 1) and empirical (Papers 2 and 3) standpoint, the findings presented in the dissertation suggest that the revised DIGCOMP framework functions adequately as a theoretical foundation and sufficiently delineates what ICT literacy should constitute from an educational viewpoint. It must be noted that this view is supported by the understanding that the ICT literacy frameworks may change due to technological innovations. New competences may be added to the frameworks, and former competences may be replaced or revised. However, the main assumption is that these competences are to a great extent generic and not related to the technology itself.

Since the respondents in Papers 2 and 3 were Norwegian teachers and students, the Norwegian ICT literacy curriculum has received substantial attention in the dissertation. Note that the comparisons between the international frameworks and the Norwegian ICT literacy curriculum were not made in the individual papers. However, taking a meta-perspective in the dissertation seemed important to investigate the alignment between the frameworks on which the papers are based and the local context from which data have been collected. The findings of the comparisons between the revised DIGCOMP and the Norwegian ICT literacy framework will therefore be discussed in the following.

### **5.2.3 The Norwegian ICT Literacy Curriculum**

Norway has received much attention as one of the first countries to integrate ICT literacy formally in its curriculum (Balanskat & Gertsch, 2010). ICT literacy is not implemented through a specific ICT subject, but rather by defining ICT literacy as one of five cross-curricular domains (or basic skills as they are formally labeled) to be addressed in all subjects. The comparison in section 2.2.5 between DIGCOMP and the Norwegian ICT literacy curriculum showed that these two frameworks do not correspond sufficiently.

A further look into the ICT literacy descriptors in the Norwegian framework identified that several competences and competence areas included in DIGCOMP are not part of the Norwegian curriculum. Furthermore, (as shown in section 2.2.5) the structure of the Norwegian framework is somewhat misleading. One categorisation level, which could potentially bridge the topical content (i.e., labeled as *categories* in the framework) and the *ability level* descriptors in the framework is missing (see Appendix A for an overview of the Norwegian ICT literacy framework). Moreover, the level descriptions in the Norwegian

framework do not correspond with the grades in the Norwegian educational system, and further explanations regarding this link are not provided in the framework (see Norwegian Directorate for Education and Training, 2012). Although it is stated in the Norwegian ICT literacy framework that the ambitions specified in the framework are to be included and operationalized in the subject-specific curricula, there are no explicit descriptions of how this should be done. Stated differently, the expert groups given the task to formulate the subject-specific curricula were not given a specification of *what* should be integrated into which subjects and during *which* year of schooling.

The subject curriculum descriptions of each subject domain contain brief descriptions of the basic skills (e.g., the science curriculum; Norwegian Directorate for Education and Training, 2013). Nevertheless, it has been argued that the quality of the work on basic skills and the extent to which they are integrated into the subjects vary to a large degree (NOU 2014:7, 2014). An evaluation of the Norwegian national curriculum revealed that *how* ICT literacy is understood varies across schools and classrooms, and the teachers perceive it as important only for students in the beginning of primary education, instead of skills that are continuously developed as part of their subject domain learning throughout their educational training (Aasen et al., 2012). Consequently, it seems that the formal responsibility for instructing students to attain the ICT literacy goals falls between different actors, and this lack of a clear-cut ICT literacy curriculum may also affect pre-service teachers and teacher training institutions, as the framework does not clearly put forward the requirements to the teachers (Tømte, Kårstein, & Olsen, 2013).

These findings are problematic in several ways and may have consequences for the development of the students' ICT literacy. First, ICT literacy has been pointed out as vital for young people, and it has been identified that students lack necessary ICT competences (Boyd, 2014; Helsper & Eynon, 2010; Selwyn, 2009), and it has been emphasized that students' knowledge of ICT should be developed within the educational system (Griffin et al., 2012). Second, because the responsibility to prepare and educate students with necessary ICT competences is placed at schools, with teachers as the key facilitators, an incomplete framework without any clear guidelines makes this expectation challenging. Third, even though the assessment of ICT literacy based on performance-based tests is seen as critical (Darling-Hammond & Adamson, 2010), to measure competences that students have not been (formally) introduced to at school seems perplexing. Fourth, from a research point of view, a lack of a clear national framework makes it challenging to design observation and



intervention studies which could potentially identify the factors that promote or inhibit students' learning of ICT literacy or teachers' instructional practice.

Consequently, revisions of the Norwegian ICT literacy curriculum are needed. An update of the curriculum and alignment with other national and international frameworks would be beneficial for teachers and students. There is also a need for broadening the focus of the Norwegian framework, as international trends emphasize the ubiquity and rapid technological changes which demand 21st-century skills. This view is also in line with the recommendations of a recent official Norwegian report, *The School of the Future* (NOU 2015:8, 2015). On the basis of the changes in society which point “in the direction of a society that has greater diversity, a high degree of complexity, and rapid changes” (NOU 2015:8, 2015 p. 8), the report recommends four areas of competence as the basis for renewing the content of school: subject-specific competence; competence in learning; competence in communicating, interacting, and participating; and competence in exploring and creating. These competences are to great extent in line with the 21<sup>st</sup>-century skills frameworks (Gallardo-Echenique et al., 2015; Griffin et al., 2012; Partnership for 21st Century Skills, 2012).

## 5.3 Assessment

The second theme identified for discussion across the three papers is *assessment*. As previously described, *assessment* is understood as a broad concept, including a variety of tools for obtaining information from different sources about achievement or abilities of individuals. Thus, assessments vary considerably, and the TEDDICS and LDN-ICT scales serve as examples of two different assessments. Investigation of the quality (reliability and validity) of the instruments is a core aspect of assessment and has been emphasized in the dissertation because purposive and high-quality assessments are necessary for drawing inferences based on the instruments.

In the following, the most important findings related to *assessment* are discussed under four sub-categories: assessment lags behind, reporting of the quality of tests, students' ICT literacy, and teachers' emphasis on ICT literacy.

### 5.3.1 Assessment Lags Behind

The findings of the systematic review showed that the ICT literacy tests do not measure important aspects of ICT literacy as defined in the most current frameworks. In addition to the imbalance between the theoretical view on ICT literacy as conceptualized in the frameworks and the assessment of it, it was also revealed that international comparative studies are to a great extent missing. In fact, except the ICILS, there has been no international study on ICT literacy. ICILS was conducted in 22 countries and operationalized parts of the ICT literacy framework, particularly emphasizing the competence area *Information*. Therefore, assessments that measure other aspects of ICT literacy are lacking. The LDN-ICT test adds to the field because it operationalizes additional facets of ICT literacy (i.e., *Digital Communication* and *Collaborative Problem Solving*), which have been scarcely operationalized in previous research. The findings of the LDN-ICT paper provide evidence for the validity and reliability of the test, and it could therefore be used to facilitate the assessment of students' competences, about which we know very little. Given that the competences outlined in the ICT literacy frameworks are cross-cultural and relevant for all students, larger international collaborative projects may be of extended value to develop authentic and innovative assessments of ICT literacy. In line with this view, Paper 3 serves as a feasibility study and an example of a cross-country collaboration.

In the national context, Norwegian students' ICT competences have been monitored since 2003 (Kløvstad & Kristiansen, 2004). However, these were measured by students' self-reports, which have been identified as confounding because biased results have been revealed when compared to performance-based assessments (Kaasbøll, 2012; Ross, 2006). In contrast, Australian students' ICT literacy is regularly monitored with use of performance-based tests in cycles of every three years (MCEETYA, 2007), and Australia is the only country which has such an assessment program regarding ICT literacy (see Paper 1). Although attempts have been made to measure Norwegian students' ICT literacy with performance-based tests (ITU monitor, 2009), they tend to include items with limited interactivity (e.g., multiple-choice items) and are often published as national reports, which do not contain information about the test-development processes or the validation of the tests. Hence, there is a lack of ICT literacy tests with more authentic tasks which require interactivity in the Norwegian educational system.

Although it is out of the scope of the dissertation to identify reasons for why assessment is lagging behind, some possible hypotheses about plausible mechanisms can be

posed: (1) Some competences might be easier to measure than others, or there has been a longer tradition of measuring them. For instance, information literacy, which includes competences such as searching for and evaluating information, has long traditions in library research (Bawden, 2001). These assessments have been further developed following the digitalization of information and transformed into information literacy, and they are in some frameworks considered as a skillset of 21st-century skills (Griffin et al., 2012). (2) The international frameworks like DIGCOMP and 21st-century skills are visionary documents and therefore more broad, whereas national curricula tend to be more conservative. (3) The development of new types and modes of assessment lags behind because it requires time to develop and validate such assessments. (4) The cost related to the development of more authentic performance-based assessments has been identified as one of the key barriers, especially compared to traditional multiple-choice tests (Jamieson, 2005; Lennon et al., 2003; O'Neil, Baker, & Perez, 2016). As such, development of ICT literacy assessment through collaborative efforts may be a more cost-efficient approach.

### **5.3.2 Reporting of the Quality of Tests**

The reliability and validity evidence stood out as an important aspect for the evaluation of the quality of assessments. For being able to draw inferences based on the test results, it is critical to investigate and ensure that the scales measure the intended construct, that they are fair across different groups, and that the development process is transparent.

There has been a transition over the past few years from paper-and-pencil tests toward computer-based assessments (Csapó, Ainley, Bennett, Latour, & Law, 2012). This means that the way students are assessed is different, and the tests have features that may affect students' performance. The advantages of this shift include automated scoring and feedback, optimized costs of test administration, and presentation of items via multimedia features (e.g., Kuo & Wu, 2013; Stacey & Williams, 2013). On the other hand, they often involve more complex skills (e.g., interactive tasks which involve solving dynamic problems) than those required in paper-and-pencil tests (Funke, 2010; Greiff, Kretzschmar, Müller, Spinath, & Martin, 2014). With regard to validity issues, research has pointed out that computer-based assessments have the potential to provide better evidence for construct validity, by collecting data describing how students approach tasks when they are involved in solving real-world problems (Chua, 2012; Messick, 1995). This is particularly important in the context of ICT literacy assessment, in which digital tools are a critical part because what we are attempting to measure is related

to technology. Items which mimic “real-world tasks” are considered to be more authentic and valid than traditional item designs such as multiple-choice tasks (Wirth, 2008).

As shown in the systematic review, the ICT literacy assessments vary to a large degree, both regarding content covered and the quality with which they were reported. For investigating the latter, 15 indicators (of reliability and validity; see Paper 1) were identified as essential. Given that the reports from these assessments gave reference to a spectrum of validity literature (or none), it was not considered useful to impose one particular theoretical position in reviewing their quality. Accordingly, the indicators finally chosen for evaluating the quality with which the validation of the assessments were reported were mostly identified and selected across dominant theories and widely used standards for validation. This finding in itself suggests that there is a need for updated validation frameworks which take into account the transition to computer-based tests and align more with the 21st-century education. Newton and Shaw (2014) have made an attempt to address this and suggested a new framework for 21st-century evaluation. However, the framework is fairly general, which makes it challenging to apply it for appraising the reported quality of tests. Besides, the novelty of the framework and how it addresses the challenges of 21st-century assessment are not clear. The framework does, however, add to the literature of general validity frameworks and standards, but it lacks the descriptors which could be applied to assessments in 21st-century education, which involves much more than the tasks presented in traditional formats. ICT literacy assessments typically involve digital tools and different levels of interactivity in task presentation to solve the task and to give a response.

To meet some of the challenges described, two indicators were developed to address some of the peculiarities of ICT literacy tests. First, as shown in section 2.3, for capturing the authenticity and innovativeness of the task/item designs, four categories were developed on the basis of frameworks for constructing innovative item types (Parshall, Harmes, Davey, & Pashley, 2010; Scalise & Gifford, 2006). The indicator was used to inspect the task/item design according to how the information (e.g., dynamic, static) was presented, what the test-taker had to do to find the right answer, and how the response was given (e.g., multiple-choice, performance). Second, an indicator labeled *explicit content* was added to inquire the level of information provided about the content of the test. This indicator was added because the reporting of what was measured in the tests varied, and in some cases it was briefly reported that the test measured, for instance, ICT literacy (the framework level) or information and communication (the level of competence area), without further detailed

explanation. Given the multifaceted frameworks of ICT literacy, the extent to which information about what the test measures is provided affects the trustworthiness of the test.

The systematic review found that few assessments are accompanied with documentation of qualities regarded as essential in order to be able to establish arguments for why interpretations of the scores should be trusted. A deficient reporting of the quality of the tests has several implications and is a central concern for the trustworthiness of the tests and the credibility of the inferences drawn by use of such tests (Wilson & Sloane, 2000). It may further affect the replicability of the tests and thus the potential to make use of the instruments in future studies and samples (Duncan, Engel, Claessens, & Dowsett, 2014). Moreover, an adequate reporting of the tests is essential because it allows test developers and researchers insights into potential available tests.

The indicators of reliability and validity were further used as a template for reporting the TEDDICS and the LDN-ICT studies, which will be described in the following section. The goal was to investigate their feasibility in being applied to similar empirical studies to further inspect the validity of the indicators.

### ***Two Examples of Validation Studies***

The aim was to report all 15 indicators set in the systematic review in the validation papers of TEDDICS and LDN-ICT. This proved to be a challenging task, particularly because it is not entirely straightforward to achieve this in *one* publication given the limitations in academic journals. However, the TEDDICS and LDN-ICT scales were developed within larger international projects, and pilot studies of both had already been conducted and reported. Collectively, the previous reporting of the scales and the two papers in the dissertation successfully demonstrated high levels of transparency in the reporting of what was measured and how (i.e., especially relevant for the LDN-ICT test because it is a performance-based test). Hence, the reliability and validity indicators previously mentioned were extensively reported (TEDDICS reported 12 indicators; LDN-ICT reported 14 indicators). The few indicators not reported for these two instruments were mainly seen as irrelevant given the purpose of the instruments. For instance, the indicator *Performance Levels* (i.e., whether performance levels or standards were developed with reference to an acknowledged methodology) did not fit to the scope of either of the two papers. In retrospect, it is relevant to note that the paper reporting the LDN-ICT test should and probably would have been identified and included in the systematic review. It may therefore be concluded that

Paper 3 serves as an example of a validation study which successfully manages to report vital aspects of the indicators.

In addition to providing knowledge about the technical aspects of the assessments, the analyses of Paper 2 and 3 revealed further insights into characteristic of the respondents, as well as relations between the scores on the scale being measured and attributes such as self-efficacy, perceived usefulness, gender, and socio-economic background (which were investigated primarily for examining the external validity of the scales). In the following sections, some of these findings are discussed in a broader context to provide information about how the results of the assessments can be interpreted and used.

### **5.3.3 Students' ICT Literacy**

The positive correlations between the students' test scores, SES, and academic aspirations support previous results which identified students' economic, social, and cultural status as strong predictors of ICT literacy achievement (Aesaert et al., 2014; Claro et al., 2012; Claro, Cabello, San Martin, & Nussbaum, 2015; Kim et al., 2014). Because academic aspirations are also related to students' home backgrounds (Fraillon et al., 2014; Hatlevik & Gudmundsdottir, 2013), these findings together point toward the digital divide, which is described as inequality related to ICT literacy and use, and reflect the differences in students' knowledge, skills, and abilities related to their socio-economic background (van Dijk, 2006). In a recent study, the effects of students' SES on their performance on a digital skills test was compared with their performance on standardized language and mathematics tests (Claro et al., 2015). The results showed that the effect of SES was even stronger for ICT literacy achievement. In particular, these results points towards the misconception that access to technology and digital environments provides equal opportunities to younger generations. One explanation may be that solving problems is more demanding in digital contexts (DiMaggio, Hargittai, Celeste, & Shafer, 2004; Hargittai, 2008). Furthermore, it has been identified that instead of leveling ICT literacy, technologies tend to amplify existing inequalities (Toyama, 2011). However, the digital divide pointed out in previous literature seems to be prevalent also with regard to students' learning in digital networks (e.g., *Information Search, Retrieval, and Evaluation; Communication; and Collaborative problem-solving*). The effects of SES on ICT achievement accentuates the responsibility of schools as important institutions for providing students with equal opportunities and preparing them with the knowledge expected for the 21st century—and this is equally relevant in a wealthy

country like Norway where access to technology in both students' homes and schools is more or less secured for all.

No significant gender differences were indicated in achievement on the LDN-ICT test. These findings are in line with some previous studies that could not identify differences across gender (Durndell & Haag, 2007; Pamuk & Peker, 2009). However, research on gender differences in ICT literacy has reported contradicting results. Some studies have shown that female students perform higher than male students (Fraillon et al., 2014), whereas others have identified differences in favor of boys (Kuhlemeier & Hemker, 2007; Li & Kirkup, 2007). Some researchers have pointed out that the gender gap with regard to ICT is decreasing (Schumacher & Morahan-Martin, 2001). Another explanation might be that the differences depend on what is being measured in the tests. For instance, studies that focus on certain competence areas lead to more detailed results than general measures. It has been shown that girls report higher competences in online communication, whereas boys show higher competences in technical operational skills (Bunz, Curry, & Voon, 2007). Hence, research on gender effects draws no clear picture on the significance of differences. Consequently, further research that takes into account the different aspects of ICT literacy and gender differences is needed. From a validation perspective, this reminds us that if the tools are to be used for reporting such inequalities, bias in the instruments themselves, such as differential item functioning, needs to be investigated and reported.

Another finding was the positive correlations between students' achievement on the LDN-ICT test and their self-efficacy, which indicates that students who believe in their own competence of using ICT for different school-related tasks also perform higher on the test. One implication may be that increased training in use of ICT for school-related tasks may increase students' ICT literacy. And since ICT competence descriptions are embedded in the subject domains (at least in the Norwegian educational system), it may affect the students' achievement in the subject domain. However, this finding calls for increased focus on the development of students' performed and perceived ICT literacy. Hence, activities enhancing students' beliefs in their own competences should deliberately be included, in addition to those aiming at developing their performance. For this purpose, teachers who emphasize the development of students' ICT competences are important. Hence, the teacher perspective will be discussed in the following with reference to findings of Paper 3.



### 5.3.4 Teachers' Emphasis on ICT Literacy

The TEDDICS construct aims at investigating to what extent teachers emphasize developing their students' ICT competences. TEDDICS differs from measures traditionally used in research on teachers' integration of ICT in classrooms, which most often comprise of teachers' use of ICT and their beliefs and attitudes towards ICT (e.g., perceived usefulness, ICT self-efficacy; Chien et al., 2014; Compeau, Higgins, & Huff, 1999). In particular, the ICT use construct, which refers to how often teachers use ICT (quantity) for different purposes, has often been considered a proxy for the extent to which teachers integrate ICT in their classroom activities. Note that these measures most often address *only* the teachers and do not refer to observations of actual teaching, what is taught, or reports from the students. There are also other research traditions which are concerned with teachers' integration of ICT in their educational practice. One focuses on the *technology acceptance model* (TAM) (Teo, 2011). This model aims to identify the level of teachers' acceptance of ICT and hence predict to what extent they are prepared to teach with ICT. The model includes, for instance, constructs on ICT anxiety, perceived ease of use, in addition to those mentioned above, and it is their relations that are examined (Ifenthaler & Schweinbenz, 2013; Wong, Teo, & Russo, 2012).

Another line of research which is concerned with educational practices and technology is *Technological Pedagogical Content Knowledge* (TPACK) (Koehler & Mishra, 2009), which is a framework considering how technology intersects with pedagogical and content knowledge. TPACK is derived from the need to better align teachers' preparation in the integration of technology with pedagogical approaches and curriculum (Sang, Valcke, van Braak, & Tondeur, 2010; Valtonen et al., 2015). Moreover, it has been argued that teachers must be competent in and able to integrate all three types of knowledge in order to successfully integrate ICT in their teaching practices (Schmidt et al., 2009).

In addition to the research traditions described, there are also a number of studies that inspected the implementation of ICT in the classroom by, for instance, observation and/or intervention studies. Yet little research is concerned with how teachers take on the task of developing students' ICT literacy as described in the curriculum. The TEDDICS measure is one approach toward using self-reports as a driver for aligning teachers' classroom practice with the development of students' ICT literacy. In particular, because a direct correspondence between TEDDICS and a common frame of ICT literacy standards could be established, the scale provides useful information on how teachers take on the responsibility of promoting these competences in their classroom practices. Thus, the TEDDICS scale may be added to



the TAM and/or TPACK models, and further investigations may reveal useful insights into how teachers could be prepared to teach ICT literacy.

The external validity of the TEDDICS scale was established by investigating the relations between the TEDDICS scale and constructs, such as teachers' use of ICT, perceived usefulness, and computer self-efficacy, showing particularly high correlations between teachers' computer self-efficacy and their emphasis on developing students' digital skills. These relations advocate that if teachers are expected to instruct students in order to improve their ICT literacy, self-confidence in their own ICT competences may be beneficial in order to meet the instructional expectations (Niederhauser and Perkmen 2010). Henceforth, teachers who do not see themselves as competent in these matters are less likely to emphasize the development of students' skills. Moreover, teachers who believe that using technology would enhance their job performance (i.e., perceived usefulness; Teo, 2011) put more emphasis on developing their students' ICT literacy. These results point toward the importance of teachers' self-beliefs with regard to the integration of ICT in their classroom activities. Thus, the development of teacher training programs may take these perspectives into account and include hands-on teaching experiences in order to strengthen the teachers' ICT self-efficacy and perceived usefulness of ICT (Hennessy, Ruthven, & Brindley, 2005; Tschannen-Moran & Woolfolk Hoy, 2007). It has been identified that newly graduated teachers do not feel digitally competent, and teacher training institutions lack programs for preparing teachers for instructional practices which integrate ICT (Tømte et al., 2013). Hence, the results of TEDDICS could also be useful for teacher training institutions in preparing teachers to utilize ICT in a sound pedagogical and didactical way to meet the requirements of preparing students for the knowledge society.

# 6 Contributions, Limitations, and Closing remarks

In this chapter, the main contributions of the study are described, the limitations and future directions are addressed, and finally closing remarks are provided.

## 6.1 Contributions

The work presented in the dissertation contributes to the knowledge on ICT literacy assessment in education. In the following section, the theoretical, empirical, and methodological contributions of the dissertation are discussed. Moreover, the implications for relevant stakeholders (e.g., researchers, policy makers, educators, teachers) are addressed.

### 6.1.1 Theoretical Contribution

The theoretical relevance of the dissertation is reflected in the attempts to develop the “tools” to assist a fair appraisal of the studies included in the systematic review. First, the structure and content of the DIGCOMP framework were revised to better align with the competences measured in the tests, and the revisions were made on the basis of descriptions of competences or competence areas which clearly showed misfit (see Paper 1 for more detailed descriptions). The revised DIGCOMP has further been applied as a research lens for inspecting the alignment with the international frameworks (i.e., CIL and 21st-century skills) and the Norwegian ICT literacy curriculum, and it has functioned satisfactorily as a reference framework in the dissertation. Hence, the revised DIGCOMP adds to the research literature as it has been proven to have the necessary breadth and depth to cover competence areas and levels, and it could therefore act as a blueprint in future studies of students’ learning of ICT literacy. Thus, the revised DIGCOMP framework could potentially be applied to other sample populations such as teacher education, school curriculum, and students. Intrinsically, one theoretical contribution is the illustration of how researchers can employ such a model to critically reflect on ICT literacy across studies of *intended*, *implemented*, and *attained* curriculum. Thus, the framework can be regarded as a useful model to bridge the gap between closely related concepts and frameworks across different contexts.

The second theoretical contribution lies in the indicators developed for the investigation of the reported quality of the tests (reliability and validity argument), which were identified and applied in the absence of a unified and generic framework. This list of indicators may lay the ground for further development of a pragmatic and practically useful framework for reporting the quality of assessments, which in turn could lead to improved and more consistent reporting of ICT literacy tests in particular and computer-based assessments in general. However, it should not be forgotten that validation studies depend on building arguments for specific uses or interpretations of data; hence, they need to be contextualized by the purpose. The third theoretical contribution comprises the four categories which were identified for appraising the innovativeness of task/item designs in the tests (see Paper 1). These were identified and coupled because a lack of a framework for appraising the diverse content of the ICT literacy tests was apparent. However, detailed taxonomies for development of innovative task/item designs in digital environments exist (Parshall et al., 2010; Scalise & Gifford, 2006), but applicable categories for evaluating the task or item design could not be identified. Therefore, the four categories may contribute to the field and function as a template which could be further developed as a basis for evaluation of task/item designs of computer-based tests.

### **6.1.2 Empirical Contribution**

#### ***Paper 1***

The systematic review synthesized “all” research on ICT literacy assessment and provided up-to-date knowledge about what is being assessed, how it is being assessed, and the reported quality of the assessments. Furthermore, by showing that the tests address many of the same issues despite the concepts used, my hope is that this work may lay groundwork for further research to communicate and strive to align with existing frameworks and concepts. Thus, this dissertation can be considered a contribution to the literature and the field of ICT literacy, test development, and assessment, as it attempts to integrate theoretical and empirical research to investigate the very competences being addressed in several ways and the quality of the assessments.

## ***Paper 2***

This dissertation adds to the research literature with the validation of the TEDDICS scale, which is unique in the sense that it adds a new perspective to research on teachers' integration of ICT. TEDDICS comprises a qualitative component of ICT use and can be used as the link between teachers' classroom practices and students' ICT literacy. The evidence for the scale's generalizability and external validity provides future researchers with an instrument to measure the emphasis teachers put on developing their students' ICT competences in a reliable and valid way with large-scale samples. Moreover, the evidence for the three factors of the measure provides more detailed information about the teachers' classroom practice than a unidimensional model. The positive relations between TEDDICS and teachers' use of ICT, computer self-efficacy, and perceived usefulness facilitate its integration in models such as the TAM and TPACK. Moreover, the alignment with the ICT literacy frameworks (i.e., the ICT literacy goals the students are required to attain) makes it increasingly relevant in studies on students' ICT literacy assessment to further investigate the relations between students' ICT competence and how teachers teach the curricular goals.

The knowledge about the relations between TEDDICS and teachers' computer self-efficacy, perceived usefulness of ICT and ICT use, age and gender contributes to the knowledge pool on teachers' ICT integration in educational practices and could be used to design teacher training programs.

## ***Paper 3***

The LDN-ICT study contributes to the research literature in several ways. First, the further development and validation of the test prove that the test can be used across countries, and the competences we are dealing with are cross-national (particularly common in Western societies). Second, it adds to the research literature by providing evidence for the feasibility of measuring competences described in the ICT literacy frameworks, which have been scarcely measured. The content of the test mimics to a large degree real-world tasks by including interactivity, open access to the Internet, and freely available tools for communication and collaboration. The LDN-ICT study comprises a considerable contribution to research on test development of 21st-century skills. Moreover, the study provides data on factors that are related to students' achievement on the test, such as self-efficacy, academic aspirations, and SES. Last, the empirical evidence for the alignment with the underlying framework (i.e., the

four dimensions) contributes to the research literature, which has been concerned with the need for empirical support for the dimensionality of the frameworks (Care, Scoular, & Griffin, 2016; Hesse, Care, Buder, Sassenberg, & Griffin, 2015). Consequently, the future use of the LDN-ICT test may reveal the strengths and weaknesses of students' competences related to each of the four dimensions and consequently inform classroom instruction and the development of a pedagogical continuum for planning and assessing ICT literacy (Voogt, Knezek, Cox, Knezek, & ten Brummelhuis, 2011).

No prior empirical study in this field has investigated the external validity of an assessment which includes students' synchronous communication and collaborative problem solving. Thus, this dissertation can be considered an original empirical contribution to the literature in the field of ICT literacy, test development, and assessment.

### **6.1.3 Methodological Contribution**

#### ***Paper 1***

The systematic review methodology by itself is by no means new, but the application of it to the field of ICT literacy assessment is rather rare. There are large numbers of systematic reviews that deal with the theoretical concepts or frameworks and synthesize them (Dede, 2009; Ferrari, 2012; Gallardo-Echenique et al., 2015; Sefton-Green, Nixon, & Erstad, 2009; Voogt & Roblin, 2012). However, no systematic review has been concerned with assessment studies or what the tests actually measure under the name of the broad concepts (e.g., ICT literacy, digital competence, digital skills). Hence, the systematic review approach taken in this dissertation contributes to the field. Another methodological implication regards the use of innovative research designs and methods for collecting data. For example, the use of LinkedIn and ResearchGate for detecting ICT literacy tests was helpful, and tests that otherwise would not have been captured by the database searches were identified. This can perhaps be considered an innovative method and thus holds methodological implications for collecting data in systematic reviews for identifying *all* studies, particularly for reviews where the phenomenon under study is likely also to be reported in so-called grey literature.

#### ***Paper 2***

In Paper 2, ESEM was used in comparison with the CFA approach to investigate the factorial structure of the TEDDICS scale. *A priori* knowledge about the underlying

framework directed the study design and was proven to be useful. In particular, comparing two relatively well-known approaches which are based on distinct assumptions about the theoretical underpinnings of the study revealed the strengths and weaknesses of the construct under investigation. Moreover, researchers have emphasized that the investigation of measurement invariance is vital for making comparisons across groups (Millsap, 2011). Moreover, comparisons across groups without establishing measurement invariance may result in biased estimates and lead to erroneous interpretations (Teo, 2015). In Paper 2, evidence for the measurement invariance was established before the data were further analysed to make comparisons across groups (e.g., gender, main subjects).

### *Paper 3*

In Paper 3, IRT was used to inspect the validity of the LDN-ICT test. IRT has many obvious advantages for both test development and reporting (see section 3.3.3), and it is therefore generally regarded as the preferable “toolkit” for developing new tests. In particular, unidimensional and multi-dimensional Rasch models were compared to investigate the dimensionality of the test aligned with the underlying framework. The multi-dimensional Rasch model has been used rarely in investigations of validity of multidimensional constructs of ICT literacy. Hence, Paper 3 adds to the research literature by providing evidence for its applicability in this field. This dissertation, however, might be the first to collect data on a novel ICT literacy test which involves multimedia, uses shared software for collaboration, and requires synchronous communication and collaborative problem solving. In turn, this could also be considered an empirical contribution as well as a methodological one.

Paper 2 and 3 used quantitative methods; therefore, one might get the impression that this dissertation is merely quantitative. Acknowledging this, it is important to emphasize that several approaches which are more qualitative in nature were used to prepare and further develop the test validated in Paper 3 (e.g., think-aloud protocols, interviews). Hence, the research design of Paper 3 contributes to the knowledge regarding test development. Especially in the field of ICT literacy, where it is costly to develop tests which mimic real-world tasks, this approach may contribute as an example of refining tests in collaborative projects.

In addition, the transparency of methods used in the dissertation can be considered a methodological contribution. Throughout the research process, the intention has been to promote transparency and to encourage researchers in the field to replicate. In brief,

methodological transparency was achieved in the following ways: Throughout the three research studies, the research design, methods for data collection, and the process of data analysis were thoroughly described and scrutinized along with potential validity threats such as research bias. For instance, inclusion/exclusion criteria, database searches, and article hits were well documented and explained in Paper 1. Moreover, in Papers 2 and 3, methods were thoroughly described in every step of the analytical process.

## 6.2 Limitations and Future Directions

The importance of ICT literacy assessment in relation with the *intended*, *implemented*, and *attained* curriculum has been explored in this dissertation. Although the dissertation aims to shed light on several relevant aspects of the field, these are only small fragments of potentially fruitful approaches, and it has several limitations. The main limitations of this work are described in the individual papers; yet some prevailing limitations with suggestions for future research are described in the following section.

### 6.2.1 Limitations in the Three Papers

#### *Paper 1*

In the systematic review, a number of indicators were developed as an attempt to create a tool for investigating the reported quality of the tests to synthesize and inform the research field. These are by no means exhaustive and should be investigated further. Moreover, the level of detail of some of the indicators had to be kept at a minimum to make them applicable and fair across the studies appraised. For instance, the indicator *Qualitative Information* (the test development included the collection of qualitative information to support arguments about what the test measures) was evaluated due to whether it was included or not, and not with regard to the extent or the methods used. This calls for further research, which may employ and further develop such frameworks to investigate the usefulness of the indicators in different computer-based assessments.

#### *Paper 2*

Relatively high correlations between the three factors of TEDDICS were identified, suggesting that the differentiation of TEDDICS is not clearly evident in this sample of Norwegian teachers. One explanation may be that teachers who emphasize the development

of students' ICT skills in one of the three hypothesized factors may put emphasis on the other factors to the same extent. Another explanation may lie in the fact that each of the factors contained only a limited number of items, which may not necessarily provide enough indicators in order to distinguish among the three TEDDICS factors. Therefore, it is suggested that future research further develop and empirically investigate extended measures of the TEDDICS construct. Besides, the TEDDICS construct should be further developed to comprise other facets of the ICT literacy framework, such as *Collaborative Problem Solving*, *Safety*, and *Communication within Digital Environments* to identify the extent to which teachers emphasize developing these competences in their classrooms.

TEDDICS was investigated on the basis of self-reports, which reflect teachers' perceptions of emphasizing students' digital skills and may therefore differ from actual classroom practice. Moreover, information regarding teachers' understanding of *what* ICT literacy is and *how* they think it should be taught could be useful; hence, future research may explore these views through interviews and/or observational data.

### ***Paper 3***

The correlations between the four dimensions of the LDN-ICT test were moderate to high. Also, for the LDN-ICT test, further development of the test is needed, and more items should be added to the two dimensions which included few items (i.e., *Social Capital* and *Intellectual Capital*). Moreover, one possible explanation for the high correlations between the factors in the two scales (TEDDICS and LDN-ICT) may be due to the fact that the competences outlined in the ICT literacy frameworks are closely related and may be better understood as processes and not clearly distinct competences. Furthermore, they often appear in a sequence, which may make it difficult to disentangle them. For instance, in many tests, one of the first tasks was related to searching for and finding information; second, the students' evaluated the information; and they subsequently created content based on the information, which was communicated to a given audience in the end. Thus, future research may investigate the alignment between the theoretical frameworks and the operationalization of them while taking these perspectives in account.

The TEDDICS and the LDN-ICT scales have operationalized some corresponding competences of the ICT literacy framework. However, the two instruments were not part of the same study; consequently, no further conclusions regarding the student-teacher interaction can be drawn. Ideally, it would be valuable to have studies which integrate the teacher and



student perspectives aligned with the framework. The ICILS had the potential to investigate these dimensions combined since the project also gauged students' ICT literacy across 22 participating educational systems (Fraillon et al., 2014). Nevertheless, due to the design of ICILS, it was not feasible to link the teachers to students or classes of students in ICILS 2013 (Fraillon et al., 2014). Hence, it was not possible to investigate the impact of TEDDICS on students' actual ICT competences. This information would be desirable in order to gain thorough knowledge, and it may be achieved by integrating TEDDICS and LDN-ICT in a future study.

## 6.3 Closing Remarks

ICT literacy has been emphasized in the dissertation as an educational outcome whose core value makes it important from several perspectives. Hence, assessment of these competences is highlighted as a help to monitor and pinpoint the development and status of ICT literacy in classrooms. The dissertation has emphasized the importance of developing authentic tests with real-world tasks, starting collaborative projects with regard to test development and research, and ensuring adequate reporting of validity and reliability of the tests.

I argue that an integrated view—one which takes into account the *intended*, *implemented*, and *attained* curriculum—should be taken to identify the strengths and shortcomings of ICT literacy in teaching and learning and to foster the development of ICT literacy. Therefore, the studies described and the results discussed are there to help students, teachers, and policy makers to grapple more easily with ICT literacy and to give them anchors relevant to the educational system.

# References

- Aasen, P., Møller, J., Rye, E., Ottesen, E., Prøitz, T. S., & Hertzberg, F. (2012). *Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen* [The ‘Kunnskapsløftet’ (Knowledge promotion) reform – The role of administration and institutions in the implementation of the reform]. Oslo: NIFU and UiO.
- Abercrombie, N., Hill, S., & Turner, B. S. (1984). *Dictionary of sociology*. London: Penguin Group.
- Adams, R. J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1–23.
- Aesaert, K. (2015). Identification and assessment of digital competences in primary education (Dissertation). University of Ghent.
- Aesaert, K., van Nijlen, D., Vanderlinde, R., & van Braak, J. (2014). Direct measures of digital information processing and communication skills in primary education: Using item response theory for the development and validation of an ICT competence scale. *Computers & Education*, 76, 168–181. doi:10.1016/j.compedu.2014.03.013
- Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding*. European Commission, Joint Research Centre. Institute for Prospective Technological Studies, Seville. Retrieved from <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=4699>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Andrade, H. L., & Cizek, G. J. (2010). *Handbook of formative assessment*. New York: Routledge.
- Association for Educational Assessment–Europe (2012). *Framework of standards for educational assessment. Version 1.0*. Rome: Edizioni Nuova Cultura.
- Balanskat, A. (2009). *Study of the impact of technology in primary schools—STEPS. Synthesis report*. European Commission, Directorate General Education and Culture. Retrieved from [http://erte.dge.mec.pt/sites/default/files/Recursos/Estudos/synthesis\\_report\\_steps\\_en.pdf](http://erte.dge.mec.pt/sites/default/files/Recursos/Estudos/synthesis_report_steps_en.pdf)
- Balanskat, A., & Gertsch, C. A. (2010). *Digital Skills Working Group. Review of national curricula and assessing digital competence for students and teachers: Findings from seven countries*. Brussels: European Schoolnet.
- Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57, 218–259.
- Bingimlas, K. A. (2009). Barriers to the successful integration of ICT in teaching and learning environments: A review of the literature. *Eurasia Journal of Mathematics, Science & Technology Education*, 5(3), 235–245.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining 21<sup>st</sup>-century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st-century skills* (pp. 17–66). Netherlands: Springer.
- Black, P. (1999). Assessment, learning theories, and testing systems. In P. Murphy (Ed.), *Learners, learning, and assessment* (pp. 118–135). London: The Open University.
- Boyd, D. (2014). *It's complicated: The social lives of networked teens*. New Haven, CT: Yale University Press.
- Buckingham, D. (2007). *Beyond technology: Children's learning in the age of digital media*. Cambridge: Polity Press.
- Bunz, U., Curry, C., & Voon, W. (2007). Perceived versus actual computer-email-web fluency. *Computers in Human Behavior*, 23, 2321–2344. doi:10.1016/j.chb.2006.03.008
- Calvani, A., Cartelli, A., Fini, A., & Ranieri, M. (2008). Models and instruments for assessing digital competence at school. *Journal of E-learning and Knowledge Society*, 4, 183–193.
- Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, 58, 797–807. doi:10.1016/j.compedu.2011.10.004

- Camilli, G. (2006). Test fairness. In R.L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221–256). American Council on Education and Praeger Publishers.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Care, E., Griffin, P., Scoular, C., Awwal, N., & Zoanetti, N. (2015). Collaborative problem solving tasks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21<sup>st</sup>-century skills: Methods and approach* (pp. 85–104). Dordrecht: Springer.
- Care, E., Scoular, C., & Griffin, P. (2016). Assessment of Collaborative Problem Solving in Education Environments, *Applied Measurement in Education*, DOI: 10.1080/08957347.2016.1209204
- Chien, S.-P., Wu, H.-K., & Hsu, Y.-S. (2014). An investigation of teachers' beliefs and their use of technology-based assessments. *Computers in Human Behavior*, 31, 198–210. doi:10.1016/j.chb.2013.10.037
- Christensen, R., & Knezek, G. (2008). Self-report measures and findings for information technology attitudes and competences. In J. Voogt & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education* (pp. 349–365). New York: Springer.
- Chua, Y. P. (2012). Effects of computer-based testing on test performance and testing motivation. *Computers in Human Behavior*, 28, 1580–1586. doi:10.1016/j.chb.2012.03.020
- Claro, M., Cabello, T., San Martín, E., & Nussbaum, M. (2015). Comparing marginal effects of Chilean students' economic, social, and cultural status on digital versus reading and mathematics performance. *Computers & Education*, 82, 1–10.
- Claro, M., Preiss, D. D., San Martín, E., Jara, I., Hinostroza, J. E., Valenzuela, S., et al. (2012). Assessment of 21<sup>st</sup>-century ICT skills in Chile: Test design and results from high school level students. *Computers & Education*, 59, 1042–1053.
- Compeau, D., Higgins, C. A., & Huff, S. (1999). Social cognitive theory and individual reactions to computing technology: A longitudinal study. *MIS Quarterly*, 23, 145–158. doi:10.2307/249749
- Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technological issues for computer-based assessments. In P. Griffin, B. McGraw, & E. Care (Eds.), *Assessment and teaching of 21<sup>st</sup>-century skills* (pp. 143–230). New York: Springer Science+Business.
- Darling-Hammond, L., & Adamson, F. (2010). *Beyond basic skills: The role of performance assessment in achieving 21<sup>st</sup>-century standards of learning*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education. Retrieved from <https://scale.stanford.edu/system/files/beyond-basic-skills-role-performance-assessment-achieving-21st-century-standards-learning.pdf>
- de Ayala, R. J. (2013). The IRT tradition and its applications. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods* (vol. 1, pp. 144–169). New York: Oxford University Press.
- Dede, C. (2009). *Comparing Frameworks for "21st Century Skills"*. Boston: Harvard Graduate School of Education. Retrieved from [http://www.watertown.k12.ma.us/dept/ed\\_tech/research/pdf/ChrisDede.pdf](http://www.watertown.k12.ma.us/dept/ed_tech/research/pdf/ChrisDede.pdf)
- Dede, C. (2011). Reconceptualizing technology integration to meet the challenges of educational transformation. *Journal of Curriculum and Instruction*, 5(1), 4–16.
- DeVellis, R. (2011). *Scale development: Theory and applications* (2<sup>nd</sup> ed.). California: Sage Publications.
- DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality: From unequal access to differentiated use. In K. Necherman (Ed.), *Social inequality* (pp. 355–400). New York: Russell Sage Foundation.
- Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2014). Replication and robustness in developmental research. *Developmental Psychology*, 50(11), 2417–2425. doi:10.1037/a0037996
- Durndell, A., & Haag, Z. (2002). Computer self-efficacy, computer anxiety, attitudes towards the Internet, and reported experience with the Internet, by gender, in an East European sample. *Computers in Human Behavior*, 18, 521–535.

- Educational Testing Service. (2002). *Digital transformation: A framework for ICT literacy*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2007). *Digital transformation: A framework for ICT literacy. A report of the International ICT Literacy Panel*. Princeton, NJ: Center for Global Assessment. Retrieved from [https://www.ets.org/Media/Tests/Information\\_and\\_Communication\\_Technology\\_Literacy/ictreport.pdf](https://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf)
- Embretson, S. E., & Hershberger, S. L. (1999). *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: L. Erlbaum Associates.
- Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher*, 35(5), 14–23.
- Erstad, O. (2006). A new direction? Digital literacy, student participation, and curriculum reform in Norway. *Education and Information Technologies*, 11, 415–429. doi:10.1007/s10639-006-9008-2
- Erstad, O. (2010). *Digital kompetanse i skolen* [Digital competence in school] (2 ed.). Oslo: Universitetsforlaget.
- Eshet-Alkalai, Y. (2002). Digital literacy: A new terminology framework and its application to the design of meaningful technology-based learning environments. In P. Barker & S. Rebelsky (Eds.), *Proceedings of EdMedia: World Conference on Educational Media and Technology 2002* (pp. 493–498). Association for the Advancement of Computing in Education.
- Eshet-Alkalai, Y., & Amichai-Hamburger, Y. (2004). Experiments with digital literacy. *CyberPsychology & Behavior*, 7, 421–429.
- Fang, W. M. Y. (2012). *No pain, no gain: Opportunities and dilemmas in online learning and teaching* (Tech. Rep.). Center for Enhanced Learning and Teaching, Hong Kong University of Science and Technology. Retrieved from <http://repository.ust.hk/ir/Record/1783.1-952>
- Ferrari, A. (2012). Digital competence in practice: An analysis of frameworks. *JRC Technical Reports*. Luxembourg: European Commission, Joint Research Centre. doi:10.2791/82116
- Ferrari, A. (2013). *DIGCOMP: A framework for developing and understanding digital competence in Europe*. In Y. Punie & B. N. Brečko (Eds.), *JRC scientific and policy reports*. Luxembourg: Publications Office of the European Union. doi:10.2788/52966
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age. The IEA International Computer and Information Literacy Study, international report*. IEA: Springer Open.
- Fraillon, J., Schulz, W., & Ainley, J. (2013). *International Computer and Information Literacy Study assessment framework*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement.
- Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, 11, 133–142. doi:10.1007/s10339-009-0345-0
- Gallardo-Echenique, E. E., de Oliveira, M. J., Marqués-Molias, L., & Esteve-Mon, F. (2015). Digital competence in the knowledge society. *MERLOT Journal of Online Learning and Teaching*, 11(1). Retrieved from [jolt.merlot.org/vol11no1/Gallardo-Echenique\\_0315.pdf](http://jolt.merlot.org/vol11no1/Gallardo-Echenique_0315.pdf)
- Gilster, P. (1997). *Digital literacy*. New York: John Wiley & Sons Inc.
- Gomes, C. (2005). Integration of ICT in science teaching: A study performed in Azores, Portugal. In *Third international conference on multimedia and information and communication technologies in education* (pp. 165–168).
- Goodlad, J. I., Klein, M. F., & Tye, K. A. (1979). The domains of curriculum and their study. In J. I. Goodlad (ed.), *Curriculum inquiry: The study of curriculum practice*, pp. 43–76). New York: McGraw-Hill.
- Gough, D., Oliver, S., & Thomas, J. (2012). *An introduction to systematic reviews*. London: SAGE Publications Ltd.
- Green, H., & Hannon, C. (2007). Young people are spending their time in a space which adults find difficult to supervise or understand. Their space education for a digital generation. Retrieved from <http://dera.ioe.ac.uk/23215/1/Their%20space%20-%20web.pdf>
- Green, L. (2010). *The Internet: An introduction to new media*. Oxford: Berg Publishers.



- Green, S., Higgins, J. P. T., Alderson P., Clarke M., Mulrow, C.D., & Oxman, A.D. (2015). What is a systematic review? In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (version 5.1.0). The Cochrane Collaboration (Chapter 1). Retrieved from [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
- Greiff, S., Kretzschmar, A., Müller, J. C., Spinath, B., & Martin, R. (2014). The computer-based assessment of complex problem solving and how it is influenced by students' information and communication technology literacy. *Journal of Educational Psychology*. doi:10.1037/40035426
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013). Complex problem solving in educational contexts—Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, 105(2), 364–379. doi:10.1037/a0031856
- Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21<sup>st</sup>-century skills*. Melbourne, Australia: Springer. doi:10.1007/978-94-007-2324-5
- Hakkarainen, K., Ilomäki, L., Lipponen, L., Muukkonen, H., Rahikainen, M., Tuominen, T., ... Lehtinen, E. (2000). Students' skills and practices of using ICT: Results of a national assessment in Finland. *Computers & Education*, 34, 103–117.
- Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and practice*, 12, 535–556.
- Hargittai, E. (2008). The digital reproduction of inequality. In D. Grusky (Ed.), *Social stratification*. Boulder, CO: Westview Press.
- Hatlevik, O. E., & Christophersen, K. (2013). Digital competence at the beginning of upper secondary school: Identifying factors explaining digital inclusion. *Computers & Education*, 63, 240–247. doi:10.1016/j.compedu.2012.11.015
- Hatlevik, O. E., & Gudmundsdottir, G. B. (2013). An emerging digital divide in urban school children's information literacy: Challenging equity in the Norwegian school system. *First Monday*, 18, 4. doi:10.5210/fm.v18i4.4232
- Helsper, E. J., & Eynon, R. (2010). Digital natives: Where is the evidence? *British Educational Research Journal*, 36(3), 503–520. doi:10.1080/01411920902989227
- Hennessy, S., Ruthven, K., & Brindley, S. (2005). Teacher perspectives on integrating ICT into subject teaching: Commitment, constraints, caution, and change. *Journal of Curriculum Studies*, 37, 155–192. doi:10.1080/0022027032000276961
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem-solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21<sup>st</sup>-century skills: Educational assessment in an information age*. doi:10.1007/978-94-017-9395-7\_2
- Huggins, A. C., Ritzhaupt, A. D., & Dawson, K. (2014). Measuring information and communication technology literacy using a performance assessment: Validation of the Student Tool for Technology Literacy (ST2L). *Computers & Education*, 77, 1–12. doi:10.1016/j.compedu.2014.04.005
- Ifenthaler, D., & Schweinbenz, V. (2013). The acceptance of tablet-PCs in classroom instruction: The teachers' perspectives. *Computers in Human Behavior*, 29, 525–534. doi:10.1016/j.chb.2012.11.004
- ITU monitor. (2009). ITU monitor: Skolens digitale tilstand 2009 [ITU Monitor 2009. The schools' digital state of affairs 2009]. Oslo: Forsknings- og kompetansenettverk for IT.
- Jamieson, J. (2005). Trends in computer-based second language assessment. *Annual Review of Applied Linguistics*, 25, pp. 228–242.
- Jesson, J. K., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. Londond:SAGE Publications nLtd.
- Kaasbøll, J. (2012). *Developing digital competence: Learning, teaching, and supporting use of information technology*. Oslo: University of Oslo. Retrieved from <http://electronics.wesrch.com/pdf/EL11TZFGRQBHU>

- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: American Council on Education/Praeger Publishers.
- Kim, H.-S., Kil, H.-J., & Shin, A. (2014). An analysis of variables affecting the ICT literacy level of Korean elementary school students. *Computers & Education*, 77, 29–38. doi:10.1016/j.compedu.2014.04.009
- Kim, J., & Lee, W. (2013). Meanings of criteria and norms: Analyses and comparisons of ICT literacy competencies of middle school students. *Computers & Education*, 64, 81–94. doi:10.1016/j.compedu.2012.12.018
- Kleven, T. A. (2008). Validity and validation in qualitative and quantitative research. *Nordic Studies in Education*, 28(3).
- Kløvstad, V., & Kristiansen, T. (2004). *ITU Monitor. Skolens Digitale Tilstand 2003* [ITU Monitor. The schools' digital state of affairs 2003]. Report 1/2004. Oslo: ITU.
- Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60–70.
- Krumsvik, R. J. (2008). Situated learning and teachers' digital competence, *Education and Information Technologies*, 13(4), 279–290.
- Kuhlemeier, H., & Hemker, B. (2007). The impact of computer use at home on students' Internet skills. *Computers & Education*, 49, 460–480.
- Kuiper, E., Volman, M., & Terwel, J. (2005). The web as an information resource in K–12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research*, 75, 285–328. doi:10.3102/00346543075003285
- Kuo, C.-Y., & Wu, H.-K. (2013). Toward an integrated model for designing assessment systems: An analysis of the current status of computer-based assessments in science. *Computers & Education*, 68, 388–403. doi:10.1016/j.compedu.2013.06.002
- Lankshear, C., & Knobel, M. (2008). *Digital literacies. Concepts, policies, and practices*. New York: Peter Lang Publishing Inc.
- Law, N., Lee, Y., & Yuen, H. K. (2009). The impact of ICT in education policies on teacher practices and student outcomes in Hong Kong. In F. Scheuermann & F. Pedro (Eds.), *Assessing the effects of ICT in education—Indicators, criteria, and benchmarks for international comparisons* (pp. 143–164). Paris: OECD.
- Law, N., Yuen, H. K., Ki, W. W., Li, S. C., Lee, Y. & Chow, Y. (2000). Changing classrooms & changing schools: A study of good practices in using ICT in Hong Kong schools. Hong Kong: Centre for Information Technology in School and Teacher Education, The University of Hong Kong. Retrieved from <http://hub.hku.hk/handle/10722/54848>
- Lennon, M., Kirsch, I., von Davier, M., Wagner, M., & Yamamoto, K. (2003). *Feasibility study for the PISA ICT literacy assessment, Report to Network A*. Paris: OECD.
- Li, N., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education*, 48(2), 301–317. doi:10.1016/j.compedu.2005.01.007
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Litt, E. (2013). Measuring users' Internet skills: A review of past assessments and a look toward the future. *New Media & Society*, 15, 612–630. doi:10.1177/1461444813475424
- MacDonald, D., & Brooker, R. (1999). Assessment issues in a performance-based subject: A case study of physical education. In P. Murphy (Ed.), *Learners, learning, and assessment* (pp. 171–190). London: The Open University.
- Markauskaite, L. (2006). Toward an integrated analytical framework of information and communications technology literacy: From intended to implemented and achieved dimensions. *Information Research*, 11(3).
- Marsh, H. W., Hau, K., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics* (pp. 275–340). Mahwah, NJ: Erlbaum.
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis.

- Annual Review of Clinical Psychology*, 10, 85e110. doi:10.1146/annurev-clinpsy-032813-153700
- Marsh, H. W., Muthén, B., Asparouhov T., Lüdtke O., Robitzsch A, et al. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, 16, 439–476.
- Martin, A. (2006). Literacies for the digital age: Preview of Part 1. In A. Martin & D. Madigan (Eds.). *Digital literacies for learning* (pp. 3–26). London: Facet Publishing.
- Martin, A. (2008). Digital literacy and the “digital society.” In C. Lankshear & M. Knobel (Eds.), *Digital literacies. Concepts, policies, and practices* (pp. 151–176). New York: Peter Lang Publishing Inc.
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Los Angeles, CA: Sage.
- McLaren, J., & Zappala, G. (2002). The “digital divide” among financially disadvantaged families in Australia. *First Monday*, 7(11). Retrieved from <http://firstmonday.org/article/view/1003/924>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist*, 50, 741–749.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge.
- Ministerial Council on Education, Employment, Training, and Youth Affairs. (2005). *National Assessment Program. Information and communication technology literacy 2005, years 6 and 10: An assessment domain for ICT literacy*. Carlton: Curriculum Corporation. Retrieved from [https://www.iste.org/docs/pdfs/australia\\_ict\\_assessment.pdf?sfvrsn=2](https://www.iste.org/docs/pdfs/australia_ict_assessment.pdf?sfvrsn=2)
- Ministerial Council on Education, Employment, Training and Youth Affairs (). (2007). *National Assessment Program. ICT literacy years 6 & 10, report 2005*. Carlton South: Ministerial Council on Education, Employment, Training and Youth Affairs. Retrieved from [http://www.nap.edu.au/verve/\\_resources/2005\\_ICTL\\_Public\\_Report\\_file\\_main.pdf](http://www.nap.edu.au/verve/_resources/2005_ICTL_Public_Report_file_main.pdf)
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1, 3–62.
- Mislevy, R. J., Corrigan, S., Oranje, A., DiCerbo, K., Bauer, M. I., von Davier, A. A., John, M. (2016). Psychometrics and game-based assessment. In F. Drasgow (Ed.), *Technology and testing: Improving educational and psychological measurement* (pp. 23–48). New York: Routledge.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: Boston College, TIMSS and PIRLS International Study Center. Retrieved from [http://timssandpirls.bc.edu/timss2011/downloads/TIMSS2011\\_Frameworks.pdf](http://timssandpirls.bc.edu/timss2011/downloads/TIMSS2011_Frameworks.pdf)
- Newton, P. E., & Shaw, S. D. (2014). *Validity in educational and psychological assessment*. London/Cambridge: Sage Publications Ltd & Cambridge Assessment.
- Niederhauser, D. S., & Perkmen, S. (2010). Beyond self-efficacy: Measuring pre-service teachers' instructional technology outcome expectations. *Computers in Human Behavior*, 26, 436–442. doi: 10.1016/j.chb.2009.12.002
- North Central Regional Educational Laboratory. (2003). *21st-century skills: Literacy in the digital age*. North Central Regional Educational Laboratory (NCREL). Retrieved from <http://www.ncrel.org/engage/skills/skills.htm>
- Norwegian Directorate for Education and Training. (2012). *Framework for basic skills*. Retrieved from [http://www.udir.no/PageFiles/66463/Framework\\_FOR\\_BASIC\\_SKILLS.pdf?epslanguage=no](http://www.udir.no/PageFiles/66463/Framework_FOR_BASIC_SKILLS.pdf?epslanguage=no)
- Norwegian Directorate for Education and Training. (2013). *Natural science subject curriculum*. Retrieved from <http://www.udir.no/kl06/NAT1-03?lplang=eng>

- NOU 2015:8. (2015). *The school of the future—Renewal of subjects and competences*. Norges offentlige utredninger 2015:8. Retrieved from <https://www.regjeringen.no/contentassets/dal48fec8c4a4ab88daa8b677a700292/en-gb/pdfs/nou201520150008000engpdfs.pdf>
- NOU 2014:7. (2014). Elevenes læring i fremtidens skole. Et kunnskapsgrunnlag [Pupils' learning in the school for the future. A knowledge base]. Norges offentlige utredninger 2014:7. Retrieved from <https://www.regjeringen.no/contentassets/e22a715fa374474581a8c58288edc161/no/pdfs/nou201420140007000dddpdfs.pdf>
- Oblinger, D.G. (2004). The next generation of educational engagement. *Journal of Interactive Media in Education*, 2004(1), 10. doi:10.5334/2004-8-oblinger
- OECD. (2001). *Schooling for tomorrow. Learning to change: ICT in schools. Education and skills*. Paris: OECD, Center for Educational Research and Innovation.
- OECD. (2015). *Students, computers and learning: Making the connection*. Paris: PISA, OECD Publishing. doi:10.1787/9789264239555-en
- O'Neil, H. F., Baker, E. L., & Perez, R. S. (2016). *Using games and simulations for teaching and assessment: Key issues*. New York: Routledge and Taylor & Francis.
- Ottenbreit-Leftwich, A. T., Glazewski, K. D., Newby, T. J., & Ertmer, P. A. (2010). Teacher value beliefs associated with using technology: Addressing professional and student needs. *Computers & Education*, 55(3), 1321–1335.
- Pamuk, S., & Peker, D. (2009). Turkish pre-service science and mathematics teachers' computer related self-efficacies, attitudes, and the relationship between these variables. *Computers & Education*, 53, 454–461. doi:10.1016/j.compedu.2009.03.004
- Parshall, C. G., Harmes, J. C., Davey, T., & Pashley, P. J. (2010). Innovative items for computerized testing. In W. J. van der Linden & C. A. W. Glas (Eds.), *Elements of adaptive testing, statistics for social and behavioral sciences* (pp. 215–230). USA: Springer Science + Business Media.
- Partnership for 21st-Century Skills (P21). (2012). *Learn for the 21st century. A report and mile guide for 21<sup>st</sup>-century skills*. Partnership for 21<sup>st</sup>-Century Skills. Retrieved from [http://www.p21.org/storage/documents/P21\\_Report.pdf](http://www.p21.org/storage/documents/P21_Report.pdf)
- Passey, D., Rogers, C., Machell, J., & McHugh, G. (2004). The motivational effect of ICT on pupils. Research Report RR523. ISBN 1 84478 204 2. London: Department of Educational Research, Lancaster University. Retrieved from [http://downloads.smarttech.com/media/research/international\\_research/uk/lancaster\\_report.pdf](http://downloads.smarttech.com/media/research/international_research/uk/lancaster_report.pdf)
- Pelgrum, W. J. (2001). Obstacles to the integration of ICT in education: Results from a worldwide educational assessment. *Computers & Education*, 37(2), 163–178.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences. A practical guide*. Malden, Ma: Blackwell Publishing.
- Quellmalz, E. (2009). Assessing new technological literacies. In F. Scheuermann & F. Pedro (Eds.), *Assessing the effects of ICT in education: Indicators, criteria, and benchmarks for international comparisons* (pp. 143–164). European Union, Paris: OECD.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, 11, 1–13.
- Sang, G., Valcke, M., van Braak, J., & Tondeur, J. (2010). Student teachers' thinking processes and ICT integration: Predictors of prospective teaching behaviors with educational technology. *Computers & Education*, 54, 103–112.
- Scalise, K., & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing “intermediate constraint” questions and tasks for technology platforms. *Journal of Technology, Learning, and Assessment*, 4, 1–43.
- Scherer, R., & Siddiq, F. (2015). Revisiting teachers' computer self-efficacy: A differentiated view on gender differences. *Computers in Human Behavior*, 53, 48–57. doi:10.1016/j.chb.2015.06.038
- Scherer, R., Siddiq, F., & Teo T. (2015). Becoming more specific: Measuring and modeling teachers' perceived usefulness of ICT in the context of teaching and learning. *Computers & Education*, 88, 202–214. doi:10.1016/j.compedu.2015.05.005



- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, 42(2), 123–149.
- Schumacher, P., & Morahan-Martin, J. (2001). Gender, Internet, and computer attitudes and experiences. *Computers in Human Behavior*, 17, 95–110. doi:10.1016/S0747-5632(00)00032-7
- Sefton-Green, J., Nixon, H., & Erstad, O. (2009). Reviewing approaches and perspectives on “digital literacy.” *Pedagogies: An International Journal*, 4, 107–125. doi:10.1080/15544800902741556
- Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, 6(3), pp. 341–362. doi:10.1177/1461444804042519
- Selwyn, N. (2009). The digital native—Myth and reality. *Aslib Proceedings*, 61(4), 364–379. doi:10.1108/00012530910973776
- Senkbeil, M., Ihme, J. M., & Wittwer, J. (2013). The test of technological and information literacy (TILT) in the national educational panel study: Development, empirical testing, and evidence for validity. *Journal for Educational Research Online*, 5, 139–161.
- Shadish, W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference*. Boston: Houghton Mifflin.
- Silverman, D. (2006). *Interpreting qualitative data: Methods for analysing talk, text, and interaction* (3rd ed.). Los Angeles: Sage.
- Smith, E. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 3, 205–231.
- Smith, H. J., Higgins, S., Wall, K., & Miller, J. (2005). Interactive whiteboards: Boon or bandwagon? A critical review of the literature. *Journal of Computer-Assisted Learning*, 21(2), 91–101.
- Søby, M. (2013). Learning to Be: Developing and Understanding Digital Competence. *Nordic Journal of Digital Literacy*, 8, 134–138. Retrieved from <https://www.idunn.no/file/pdf/62414156/#page=3>
- Stacey, K., & Williams, D. (2013). Technology and assessment in mathematics. In M. A. Clements, A. Bishop, C. Keitel-Kreidt, J. Kilpatrick, & F. K.-S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 721–751). New York: Springer.
- Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer-Assisted Learning*, 21, 79–90. doi:10.1111/j.1365-2729.2005.00115.x
- Strømsø, H. I., & Bråten, I. (2014). Students’ sourcing while reading and writing from multiple web documents. *Nordic Journal of Digital Literacy*, 9, 92–111.
- Suto, I. (2013). 21st-century skills: Ancient, ubiquitous, enigmatic? *Research Matters: A Cambridge Assessment Publication*, 15, 2–8. Retrieved from <http://www.cambridgeassessment.org.uk/images/130437-21st-century-skills-ancient-ubiquitous-enigmatic-.pdf>
- Tennant, A., & Pallant, J. F. (2007). DIF matters: A practical approach to test if differential item functioning makes a difference. *Rasch Measurement Transactions*, 20, 1082–1084.
- Teo, T. (2011). Factors influencing teachers’ intention to use technology: Model development and test. *Computers & Education*, 57, 2432–2440. doi:10.1016/j.comped.2011.06.008
- Teo, T. (2015). Comparing pre-service and in-service teachers’ acceptance of technology: Assessment of measurement invariance and latent mean differences. *Computers & Education*, 83, 22–31. doi:10.1016/j.compedu.2014.11.015
- Terenzini, P. T. (1989). Assessment with open eyes: Pitfalls in studying student outcomes. *Journal of Higher Education*, 60(6), 644–664.
- Thijs, A., & van den Akker, J. (2009). *Curriculum in Development*. SLO, Enschede: Netherlands. Retrieved from [www.slo.nl/downloads/2009/curriculum-in-development.pdf](http://www.slo.nl/downloads/2009/curriculum-in-development.pdf)
- Thomas, M. L. (2011). The value of item response theory in clinical assessment: A review. *Assessment*, 18, 291–307.

- Tømte, C., Kårstein, A., & Olsen, S. D. (2013). IKT i lærerutdanningen. På vei mot profesjonsfaglig digital kompetanse? [ICT in teacher training. On the way to professional digital competence?]. Report 20/2013. Oslo: Nordisk institutt for studier av innovasjon, forskning og utdanning. ISBN: 978-82-7218-919-7
- Tondeur, J., van Braak, J., Siddiq, F., & Scherer, R. (2016). Time for a new approach to prepare future teachers for educational technology use: Its meaning and measurement. *Computers & Education*, 94, 134–150. doi:10.1016/j.compedu.2015.11.009
- Toyama, K. (2011). Technology as amplifier in international development. In *Proceedings of the 2011 iConference* (pp. 75–82). doi:10.1145/1940761.1940772
- Tschannen-Moran, M., & Woolfolk Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, 23(6), 944–956. doi:10.1016/j.tate.2006.05.003
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Valtonen, T., Kukkonen, J., Kontkanen, S., Sormunen, K., Dillon, P., & Sointu, E. (2015). The impact of authentic learning experiences with ICT on pre-service teachers' intentions to use ICT for teaching and learning. *Computers & Education*, 81, 49–58.
- Van den Akker, J. (2003) Curriculum perspectives: An introduction. In J. van den Akker, W. Kuiper and U. Hameyer (Eds.), *Curriculum landscapes and trends* (pp. 1–10). Dordrecht: Kluwer.
- Van Dijk, J. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4–5), 221–235. doi:10.1016/j.poetic.2006.05.004
- Vonkova, H., & Hrabak, J. (2015). The (in)comparability of ICT knowledge and skill self-assessments among upper secondary school students: The use of the anchoring vignette method. *Computers & Education*, 85, 191–202. doi:10.1016/j.compedu.2015.03.003
- Voogt, J. (2008). IT and curriculum processes: Dilemmas and challenges. In J. Voogt, & G. Knezek (Eds.), *International Handbook of Information Technology in Primary and Secondary Education* (pp. 117–132). New York: Springer.
- Voogt, J., Knezek, G., Cox, M., Knezek, D. and ten Brummelhuis, A. (2011). Under which conditions does ICT have a positive effect on teaching and learning? A call to action. *Journal of Computer-Assisted Learning*, 29, 4–14. doi:10.1111/j.1365-2729.2011.00453.x
- Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21<sup>st</sup>-century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, 44, 299–232.
- Vuorikari, R., Punie, Y., Carretero Gomez, S., & Van den Brande, G. (2016). DigComp 2.0: The Digital Competence Framework for Citizens. Update phase 1: The Conceptual Reference Model. Luxembourg Publication Office of the European Union. EUR 27948 EN. doi:10.2791/11517
- Warschauer, M. (2003). Social capital and access. *Universal Access in the Information Society*, 2(4), 315–330.
- Watson, D. M. (2001). Pedagogy before technology: Re-thinking the relationship between ICT and teaching. *Education and Information Technologies*, 6(4), 251–266.
- William, D. (2014). Formative assessment and contingency in the regulation of learning processes. Paper presented at *Toward a Theory of Classroom Assessment as the Regulation of Learning*, at the annual meeting of the American Educational Research Association, Philadelphia, PA, April 2014.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Wilson, M., & Scalise, K. (2015). Assessment of learning in digital networks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21<sup>st</sup>-century skills, volume 2—Methods and approaches* (pp. 57–81). Dordrecht: Springer.
- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13, 181–208.

- Wirth, J. (2008). Computer-based tests: Alternatives for test and item design. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 235–252). Cambridge: Hogrefe & Huber.
- Wong, K.-T., Teo, T., & Russo, S. (2012). Influence of gender and computer teaching efficacy on computer acceptance among Malaysian student teachers: An extended technology acceptance model. *Australasian Journal of Educational Technology*, 28, 1190–1207.
- Wu, M. L., & Adams, R. J. (2013). Properties of Rasch residual fit statistics. *Journal of Applied Measurement*, 14(4), 339–355.
- Zelman, M., Avdeeva, S., Shmis, T., Vasiliev, K., & Froumin, I. (2011). International comparison of information literacy in digital environments. Paper presented at the International Association for Educational Assessment, Manila, Philippines. Retrieved from [http://www.iaea.info/documents/paper\\_30e43f54.pdf](http://www.iaea.info/documents/paper_30e43f54.pdf)
- Zumbo, B. D., & Chan, E. K. H. (2014). *Validity and validation in social, behavioral, and health sciences*. New York: Springer. ISBN 978-3-319-07794-9.

# APPENDICES

## Appendix A. Norwegian Directorate for Education and Training 2012—Framework for Basic Skills

Digital skills as basic skills					
Field of skills	Level 1	Level 2	Level 3	Level 4	Level 5
<b>Search and process</b>	Can read hypertexts and simple interactive information. Can use picture- and icon-based navigation.	Can make simple digital searches and read and interpret information from digital sources. Can use simple digital resources and tools for information processing and learning.	Can choose and use search strategies and assess information from digital sources. Can use different digital tools and resources for information processing and learning.	Can filter, transform, and collate information from digital sources. Can use relevant search tools and master search strategies in subject-related tasks.	Can find, organize, and update digital information. Can use advanced search strategies and sources in subject-related work.
<b>Produce</b>	Can write simple texts on keyboard and produce simple composite texts. Knows simple digital use of sources and copyright rules.	Can produce digital composite texts following simple formal requirements. Can make simple use of digital sources, observing copyright rules, also in re-use and further development.	Can make digital composite texts with linked content. Can understand and use digital formal requirements in one's own texts. Can refer to digital sources and apply copyright rules.	Can produce and edit complex digital texts. Can refer to and assess digital sources in relevant subject-related situations.	Can choose and use target-group-relevant digital tools and digital formal requirements. Can administer copyright rules to one's own digital products and master digital source referencing.
<b>Communication</b>	Can use simple digital tools and media for presentation and communication.	Can use a selection of digital tools and media for presentation and communication.	Can make varied use of different digital tools and media to convey a message both in one-to-one and group communication.	Can use digital media and tools to convey a clear and detailed message for communication and documentation.	Can choose, assess, and apply digital communication tools according to different subject-related needs.
<b>Digital Judgment</b>	Can follow basic rules for digital interaction. Knows basic rules for protection of personal privacy on the Internet.	Can apply basic netiquette and knows about rules for protection of personal integrity on the Internet.	Can apply netiquette and follow rules for protection of personal integrity on the Internet and in social media.	Can use the Internet and social media efficiently and appropriately.	Can reflect ethically on and assess the Internet and social media as communication and information channels.

## Appendix B. Protocol for systematic review

### Definitions and conceptual issues

ICT literacy is a broad term, and refers to students' knowledge, skills and attitudes. No systematic review for providing an overview of ICT literacy assessments could be identified. Therefore, this review aims at contributing to fill some of the research gap. However, a large number of reviews of ICT literacy frameworks exist. These will be used to identify differences and commonalities in concepts used to address ICT literacy and the content of the frameworks.

### The overall aim of the review:

What is the state-of-the-art with regard to students' ICT literacy assessment?

In order to answer the overall aim, further research questions will be addressed:

*What are the general characteristic of the studies (country, grade, type of school, type of assessment, scoring).*

*What is being measured by the assessments (given the multidimensional frameworks)?*

*What is reported with regard to reliability and validity of the tests*

### Search strategy

**Languages:** English (but efforts will be put to try to find “all” assessments and translate if other studies are located)

**Time frame:** 1990-2014

### Sources

Electronic databases	Other sources	Journals
ERIC	Reference search (snowball)	Computers & Education
ScienceDirect	Google	Computers in Human Behavior
Web of Knowledge	Email (contacts and key authors)	
Google Scholar	LinkedIn	
ProQuest	Academia.edu	
	ResearchGate	

**Eligibility criteria**

Inclusion criteria:

Primary and secondary students (age 5-19)

Performance assessment

ICT literacy (or related concept)

Exclusion criteria:

Self-report

Assessment of subject domain knowledge with use of computer-based assessment

Second-level exclusion criteria:

Insufficient reporting of what is measured by the test

Insufficient reporting of the quality of the test

**Coding scheme**

A system for coding the included system needs to be developed to capture the relevant information of the studies to answer the research questions. Since no template which has been used in such study could be found an iterative method will be used.

**Additional notes**

There will be a need for schemes/systems/tools for appraising the studies.

# Dissertational papers

1

2

3