



**UiT** The Arctic University of Norway

Faculty of Health Sciences, Department of Psychology

University Library

# **Understanding the development of information literacy in higher education: Knowing, doing, and feeling**

Ellen Nierenberg

A dissertation for the degree of Doctor of Philosophy (PhD), September 2022



# Table of Contents

Acknowledgements .....	iv
Abbreviations .....	vii
List of papers and author contributions .....	viii
Paper 1 .....	viii
Paper 2 .....	viii
Paper 3: .....	ix
Paper 4: .....	ix
Abstract .....	x
1 Introduction .....	1
1.1 Background .....	1
1.2 Literature review .....	2
1.2.1 Information literacy .....	2
1.2.2 Transformative learning theory .....	11
1.2.3 Identity.....	14
1.2.4 Interest theory.....	18
1.2.5 Gaps in the literature .....	21
1.3 Research questions .....	22
2 Methods .....	24
2.1 Research ethics .....	24
2.2 Theoretical stance.....	25
2.3 Methodology – Mixed-methods research.....	26
2.4 Research design – Multisample Hybrid Mixed-Methods Design .....	28
3 Methods for individual studies .....	32
3.1 Methods: Paper 1.....	35
3.1.1 Information literacy knowledge test: <i>Knowing</i> .....	35
3.1.2 Source-evaluation and source-use measures: <i>Doing</i> .....	37
3.2 Methods: Paper 2.....	38
3.3 Methods: Paper 3.....	39
3.4 Methods: Paper 4.....	40
3.4.1 Quantitative methods.....	41
3.4.2 Qualitative methods.....	42
3.4.3 Focus-group interviews .....	44
3.4.4 Individual interviews.....	44

3.5	Coherence .....	45
4	Results: A short summary .....	49
4.1	Quantitative findings .....	49
4.2	Qualitative findings .....	54
4.2.1	Views of IL and of being information literate .....	54
4.2.2	Thematic analysis .....	55
4.3	Mixed methods findings .....	58
5	Discussion .....	59
5.1	Contributions to the field .....	59
5.2	Development over time .....	61
5.3	Integration and transformative learning .....	63
5.4	Relevance for IL instruction .....	65
5.5	Reflections on research process .....	66
5.6	Limitations and considerations .....	67
5.7	Future research .....	69
6	Conclusion .....	70
	References .....	71
	Paper 1 .....	1
	Paper 2 .....	1
	Paper 3 .....	1
	Paper 4 .....	1
	Appendix 1: Framework analysis .....	1
	Appendix 2: Course Certificate .....	1
	Appendix 3: NSD notification form .....	1
	Appendix 4: Consent Form .....	1
	Appendix 5: Data Management Plan .....	1
	Appendix 6: Information literacy knowledge test (TILT) .....	1
	Appendix 7: Assignment-based measures for assessing information literacy skills .....	1
	Appendix 8: Tromsø Interest Questionnaire (TRIQ) .....	1
	Appendix 9: Interest-o-meter .....	1
	Appendix 10: Interview guides .....	1
	Appendix 11: Coding of Interviews .....	1

## List of Tables

Table 1: Data Collection, Listed Chronologically, with Qualitative Strands in Grey .....	30
Table 2: Materials, Measures, Participants, Data Collection .....	33

## List of Figures

Figure 1: Pragmatism Paradigm: Conceptual Framework .....	26
Figure 2: Conceptualization of the Mixing of Methods and their Iterative Nature .....	32
Figure 3: Timeline: Development of the Tromsø Information Literacy Test (TILT) .....	35
Figure 4: Integration of Findings Across Articles .....	46
Figure 5: Mean Scores on Know, Do, and Feel (Meaningfulness Subscale) Measures, with 95% Confidence Intervals. Cross-Sectional Sample, Semesters 1 and 6 .....	50
Figure 6: Mean TILT Scores, with 95% Confidence Intervals .....	51
Figure 7: Relationships Between Know, Do, and Feel in Semesters 1 and 6: Cross-Sectional Data....	53
Figure 8: Transformative Learning: An Integration of Knowing, Doing, and Feeling, Involving Critical Reflection .....	64

# Acknowledgements

My PhD journey began in October 2018, when I moved to Tromsø and started as a PhD student in Information Literacy (IL) at UiT The Arctic University of Norway. Now, nearly four years later – and a mere 40 years after I received my first bachelor’s degree (!) – my PhD journey is approaching its end. How fortunate I’ve been to have had the opportunity to immerse myself in a research bubble, with excellent guidance, these past four years. I’m grateful to many people along the way for making this journey possible.

First of all, I’d like to thank the Library at Inland Norway University of Applied Sciences, with director Anna Løken, for offering me a four-year leave-of-absence from my position as an academic librarian to do this PhD. Without the security of a job at the end of the PhD position, I might not have had the guts to start on this journey.

I thank my family in Norway for also giving me a four-year “leave-of-absence” as I pursued my doctoral ambitions in the far north, and to my friends and family around the world for their constant support. Thank you all, from the bottom of my heart!

At UiT, I’m grateful to administrators at the University Library, the Faculty of Health Sciences, and the Department of Psychology for creating and administering the PhD position. A big thanks for your assistance along the way.

Twenty-five percent of my four-year PhD position was duty work, including teaching, developing and updating online courses, and organizing a conference. My sincerest gratitude to all of my amazing colleagues at the University Library at UiT, especially to those in:

- the work group for the online course iKomp (IL for students), including Helene Nordgaard Andreassen, Lars Figenschou, Torstein Låg, Grete Overvåg, and Mark Stenersen
- the work group for the online course on plagiarism prevention (for faculty), including Helene Nordgaard Andreassen, Marit Bull Enger, and Stian Lillejord
- the organizing committee for Creating Knowledge 2021 (international IL conference), including Helene Nordgaard Andreassen, Lars Figenschou, Øystein Lund, Mariann Løkse, and Torstein Låg

As for my wonderful colleagues at the Department of Psychology, I'm grateful to the team teaching the introductory academic writing course, including Espen Bjørkedal, Torstein Låg, Rannveig Grøm Sæle, and Morten Øvervoll. Thank you for allowing me to adjust student assignments to suit my research needs and to follow students' development over time. While at UiT, I've been a member of the research group for cognitive psychology, EPIC, led by Tove Dahl. Other members include Jon-André Dalbakk, Magnus Ingebrigtsen, Torstein Låg, Efim Nemtcan, Anett Ragó, Frode Svartdal, and Rannveig Grøm Sæle. Thank you to everyone in this group for valuable discussions, intellectual stimulation, feedback on my work, writing seminars at Sommarøy, and lots of fun!

My research would not have been possible without all of the students, both at UiT and other colleges and universities, who filled out my surveys, participated in interviews, and/or partook in think-aloud protocols. Bachelor student Mina Taaje Sønsterud made valuable contributions to the article about metacognition, and I also received time-saving help from research assistants Tina Eriksen, Amanda Lindberg, and Kristell Karlsson. I thank all of you for the time and effort you've put into this important research!

Thank you to IL experts Helena Francke, Lilja Johannessen, Grethe Moen Johansen, Mariann Løkse, and Heidi K. Olsen for your evaluations of the IL-knowledge pilot test.

And finally, I am eternally grateful to my amazing team of supervisors: main supervisor Professor Tove Irene Dahl, and co-supervisors senior academic librarian Torstein Låg and Professor Mariann Solberg. I can never thank you enough for your guidance, brilliance, engagement, constant encouragement, and your patience when my neurons were stretched beyond their limits. What a joy and a privilege it has been working with all of you, and I hope that we can continue doing research together after I'm finished with my PhD.

Mariann is the team's expert in qualitative methods. Mariann, your words and wisdom will always be with me, reminding me to define terms, clarify concepts, make logical inferences based on theory, and to substantiate claims. Working with you on interviews, coding, and the thematic analysis was like doing intellectual gymnastics! You introduced an approach rooted in educational philosophy to my quantitatively-oriented mind, opening horizons to new ways of thinking. Thank you for all of our rewarding meetings and conversations – on campus and otherwise – for your insight and thoughtfulness, and for your patience with a PhD student with little background in the humanities, such as myself.

Torstein has vast knowledge of quantitative methods, psychology, and information literacy. Torstein, thank you for encouraging me to apply for the PhD position and for being my role model as an IL teacher. You have a thorough understanding of the research process and an incredible ability to explain difficult concepts in understandable ways – even factor analyses! We’ve worked in close proximity for much of my time at UiT and have had many exchanges of ideas and fruitful discussions. Torstein, it’s impossible for me to see you without smiling (and then asking a question or two). How fortunate I’ve been to work together with you, not only with my PhD research, but also in library workgroups and in the teaching team.

And last, but not least, Tove. Tove is our team’s leader, and an expert in educational psychology and quantitative methods. What luck to have an educational psychologist as my main supervisor! Tove, I cannot thank you enough for everything that you have taught me over the past four years, both as a researcher and as a human being. You are an educational psychologist at work and in your daily life, motivating people to do their best and making learning fun (with the possible exception of discriminant function analyses!). Some of your favorite words are, “you got this.” You make everyone feel *seen*, special, competent, and cared for. You are the smartest *and* the nicest person I know. You are my role model. I remember planning one of our articles together and seeing how your mind raced. You quickly grasped the big picture, and had ideas for the research design and how this study fit together with the project as a whole. What we managed in those few hours would have taken me months on my own. Few people in this world have a skill set like yours – including research-related, creative, analytical, interpersonal, problem-solving, communicative, organizational, administrative, and social skills – and you use your outstanding abilities to make this world a better place. This you do brilliantly through your work both at the university, at Concordia Language Villages, and with various organizations. Tove, you have enriched my life and helped me to become a researcher. Without your expert guidance, I would never have gotten this far.

## **Abbreviations**

- FGI: Focus group interview
- IL: Information Literacy
- LIS: Library and Information Science
- NSD: The Norwegian Centre for Research Data
- QUAL: Qualitative
- QUAN: Quantitative
- TL: Transformative Learning
- UiT: UiT The Arctic University of Norway



# List of papers and author contributions<sup>1</sup>

## Paper 1:

Published June 2021

Nierenberg, E., Låg, T., & Dahl, T. I. (2021). Knowing and doing: The development of information literacy measures to assess knowledge and practice. *Journal of Information Literacy*, 15(2), 78–123. <https://doi.org/10.11645/15.2.2795>

Contributor roles:

- Ellen Nierenberg (main author): writing (original draft, editing, Appendix B); conceptualization and formulation of research aims; data collection, analysis, and presentation; design of methodology; data curation.
- Torstein Låg: writing (original draft, review, Appendix A); conceptualization of research aims; data analysis and presentation; design of methodology; supervision.
- Tove Irene Dahl: conceptualization of research aims; data analysis; design of methodology; supervision; project administration; writing (review).

## Paper 2:

Published November 2021

Dahl, T. I., & Nierenberg, E. (2021). Here's the TRIQ: The Tromsø Interest Questionnaire based on the four-phase model of interest development. *Frontiers in Education*, 6(402), 1-17. <https://doi.org/10.3389/feduc.2021.716543>

Contributor roles:

- Tove Irene Dahl (main author): writing (original draft, review, editing); conceptualization and design of Study 1; formulation of research goals; data collection; inferential statistical analyses; data presentation/visualization; design of methodology; creation of models; supervision; validation.

---

<sup>1</sup> Based on CRediT (Contributor Roles Taxonomy) <https://credit.niso.org/>

- Ellen Nierenberg: writing (sections of original draft, review); conceptualization and design of Study 2; data collection; analyses of descriptive data.

### **Paper 3:**

Published December 2021

Nierenberg, E., & Dahl, T. I. (2021). Is information literacy ability, and metacognition of that ability, related to interest, gender, or education level? A cross-sectional study of higher education students. *Journal of Librarianship and Information Science*.

<https://doi.org/10.1177/09610006211058907>

Contributor roles:

- Ellen Nierenberg (main author): writing (original draft, review, editing); data collection, analysis, presentation, visualization; conceptualization and formulation of research goals; design of methodology; validation; data curation.
- Tove Irene Dahl: writing (review, editing); conceptualization of research goals; data collection, analysis, presentation, visualization; design of methodology; supervision.

### **Paper 4:**

Submitted to *College & Research Libraries*, August 2022

Nierenberg, E., Solberg, M., Låg, T., & Dahl, T. I. (?). Knowing, doing, and feeling: A three-year, mixed-methods study of undergraduates' information literacy development.

Contributor roles:

- Ellen Nierenberg (main author): writing (original draft, editing); conceptualization and formulation of research aims; data collection, analyses, presentation, visualization; design of methodology; data curation.
- Mariann Solberg: writing (review); design of qualitative methods; ideas; data collection (interviews and interest-o-meter); data analysis (thematic); design of methodology; supervision.
- Torstein Låg: writing (review); data analysis; design of methodology; supervision.
- Tove Irene Dahl: writing (original draft and review); conceptualization of research aims; data analysis and presentation; supervision; project administration.

## Abstract

Information literacy (IL), the ability to recognize when information is needed and be able to find, evaluate, and effectively use this information, is frequently taught to college students. The aim of this doctoral research is to study the development of IL in undergraduates over their first three years, including their knowledge, skills, and attitudes – called *knowing*, *doing*, and *feeling* in this project – and the relationships among the three. Mixed methods are employed to better understand the totality of students' IL development, with qualitative methods added to enhance the quantitative methods during the second half of the project. Longitudinal and cross-sectional data are collected and analyzed in the quantitative study. A suite of four tools is developed to measure students' IL knowledge, skills, and attitudes. These tools are psychometrically evaluated for reliability and validity, and freely available for IL researchers and practitioners. A multiple-choice test to measure IL knowledge focuses on the three core, source-based facets of IL: finding, evaluating, and using information. Two authentic assignments are employed to assess two IL skills: the evaluation and use of information sources. IL attitudes are defined in this research in terms of interest, namely students' interest in being or becoming an information literate person, and an interest questionnaire was developed to measure phases and aspects of this interest. The questionnaire is domain-tailorable and can be used to measure other interests as well, making it a valuable contribution also in the discipline of psychology.

Aside from contributing valuable tools for measuring various aspects of IL, this exploratory research fills several gaps in the IL literature, contributing to the IL research field and to its field of practice in several ways. First, the research addresses the dimensionality of the IL construct, whether IL is one homogeneous construct or comprised of several interrelated constructs. Second, the project follows the development of IL in students over a full cycle of bachelor studies, a longer period of time than most other IL studies. Third, the research addresses relationships among students' IL knowledge, skills, and attitudes over time. Fourth, the research introduces the measurement of interest – an important motivator for learning – into the field of IL. Fifth, the research answers questions regarding the accuracy of students' estimated scores on the IL-knowledge test and factors that affect their self-awareness. Sixth, the research explores student learning through the lens of transformative learning theory to determine whether transformative learning occurs in students as a result of becoming information literate. Seventh, the project demonstrates the value of a mixed-methods methodology in the field of IL.

# 1 Introduction

## 1.1 Background

What do the war in Ukraine, protests against COVID regulations, and the storming of US Capitol have in common? All have been greatly influenced by the abundance of false information. People around the world are bombarded by huge amounts of information in their daily lives, some of which is false. In some cases, it is believed to be true by the person who disseminated it (misinformation), while in other cases, those who disseminate it know that it is not true; it is a deliberate lie, often with malicious intent (disinformation). Several studies, including Vosoughi et al. (2018) and Silverman (2016), have shown that false information spreads faster and more widely than information from trustworthy sources.

One way to curb the spread of false information is to provide people with the knowledge and skills that enable them to better discern between reliable and unreliable sources. One of the goals of information literacy instruction is to do just that. Information literacy (IL) encompasses the knowledge, skills, and attitudes needed to be able to discover, evaluate, and use information sources effectively and appropriately in order to answer questions, solve problems, create knowledge, and learn. Information-literate people are critical thinkers and lifelong learners who can find reliable information and use it to suit their needs, whether they are at college, at work, making a decision about their health, or navigating challenges in daily life.

IL is often taught in the context of higher education (HE), which is also the context of this PhD project. Norway, where this project was undertaken, has mandates for HE institutions similar to those in other parts of the world. These mandates include conducting research and disseminating its results, contributing to research-based education and to lifelong learning, facilitating employees' and students' participation in the public debate, and ensuring that Norwegian HE and research are at the forefront of international research and of developments within HE (University and University Colleges Act, 2005, § 1-3). IL is important for the achievement of many of these mandates, and IL instruction is therefore either required or strongly recommended at Norwegian HE institutions.

## **1.2 Literature review**

This doctoral dissertation is grounded in four major areas of scholarship: information literacy, transformative learning (TL), identity, and interest. In the literature review, I discuss literature and theories in these fields that have informed my research.

### **1.2.1 Information literacy**

Although IL is relevant for all disciplines, as a topic of research it is often studied by scholars in the field of library and information science (LIS). While there are few acknowledged *theories* of IL, its standards, frameworks, and models provide the foundation for much IL research and practice.

#### **Definitions, Standards, Frameworks, Models**

One of the challenges in the field of IL is the ongoing debate regarding the term itself. There is no consensus on its definition or what it encompasses.

In this section I describe frequently used definitions of IL, as well as two of the most influential standards and frameworks in the context of higher education. Both of these originated in the US, where academic librarians have worked extensively with IL instruction and assessment for many years (for an overview of other IL standards, frameworks, and models, see Appendix 1).

The term information literacy was first coined in 1974 by Paul G. Zurkowski, then president of the Information Industry Association. In a report for the National Commission of Libraries and Information Science, he wrote that:

People trained in the application of information resources to their work can be called information literates. They have learned techniques and skills for utilizing the wide range of information tools as well as primary sources in molding information solutions to their problems. The individuals in the remaining portion of the population, while literate in the sense that they can read and write ... realistically must be considered to be information illiterates. (Zurkowski, 1974, p. 6)

The definition of IL has gradually evolved due to the exponential growth of information resources, technological advancements for storing, organizing, and accessing information, and growing understandings of IL as more than a set of skills.<sup>2</sup>

In 1989, the American Library Association (1989, para. 3) created a new definition of IL: “To be information literate, a person must be able to recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information.” This definition, widely employed for over two decades, lists specific skills associated with IL. These skills were later incorporated into the *Information Literacy Competency Standards for Higher Education* ([Standards], Association of College and Research Libraries [ACRL], 2000), composed of five overarching standards with underlying performance indicators and outcomes for assessment purposes. In this definition, the information literate student: ...constructs a search strategy using appropriate commands for the information retrieval system selected (p. 9); ...creates a system for organizing information (p. 10); ...examines and compares information from various sources in order to evaluate reliability, validity, accuracy, authority, timeliness, and point of view or bias (p. 11); and ...demonstrates an understanding of intellectual property, copyright, and fair use of copyrighted material (p. 14; ACRL, 2000). The Standards were extensively employed in IL instruction and assessment in HE, in the US and internationally, where academic librarians and their partners designed learning outcomes for integrating its enumerated skills into HE curricula. Some considered this a “checklist approach” to IL instruction.

Many scholars and practitioners of IL, however, gradually began to express dissent over the Standard’s checklist approach to learning, as the realm of information was becoming increasingly complex. For example, new modes of publishing – such as blogs and Wikipedia – blurred the boundaries between information producers and consumers, making it more difficult to evaluate the reliability of information and to use it ethically. The ACRL (2015) realized that the perception of IL as a mere checklist of skills no longer harmonized with the complexities of this continuously changing information environment. They recommended therefore to focus attention on “foundational ideas” of IL rather than on rapidly shifting standards and outcomes (ACRL, 2015, p. 7). As the realm of information evolved, so did the

---

<sup>2</sup> For a thorough investigation of the evolution of the definition of IL, see Sample (2020).

responsibilities of HE faculty, students, and librarians. Teaching faculty now had a “greater responsibility in designing curricula and assignments that foster enhanced engagement with the core ideas about information and scholarship within their disciplines” (ACRL, 2015, p. 7). Students’ roles in creating new knowledge, and librarian collaborations with faculty on IL curricula and instruction, had both increased (ACRL, 2015, p. 7). As the core ideas of IL evolved, so did their complexity. After much debate (described in Sample, 2020), ACRL’s Standards were therefore rescinded in 2016 and replaced with their *Framework for Information Literacy for Higher Education* ([Framework], ACRL, 2015).

This modern conception of IL is reflected in the ACRL’s revised definition in the Framework:

Information literacy is the set of integrated abilities encompassing the reflective discovery of information, the understanding of how information is produced and valued, and the use of information in creating new knowledge and participating ethically in communities of learning. (ACRL, 2015, p. 3)

In contrast to listing IL skills as the Standards did, the ACRL Framework is comprised of six interrelated core concepts (*frames*), with underlying practical competencies (*knowledge practices*) and mindsets (*dispositions*). The Framework’s six frames are, in alphabetical order: Authority is constructed and contextual; Information creation as a process; Information has value; Research as inquiry; Scholarship as conversation; and Searching as strategic exploration (ACRL, 2015). These frames are partly informed by theories of *threshold concepts* – concepts that can generate transformation in learners by enabling the formation of new perspectives (Meyer et al., 2010). An advantage with the Framework is that it offers more flexible options for curricular development and implementation than the prescribed skills and outcomes of the Standards. In addition, the Framework represents a more holistic view and conceptual understanding of information, research, and scholarship, that better enables transformation in learners (ACRL, 2015, p. 9, Note 3).

Many librarians, however, argued that the Standards should not be discarded, but rather updated to reflect new concepts in the Framework. In an open letter to the ACRL (Berg, 2015), a group of librarians from New Jersey justified this standpoint with several arguments, including:

- The Standards are easily understood, and its outcomes-based competencies are easily employed as learning outcomes in curricula.

- The Standards express uniform goals and acceptable levels of achievement.
- The Standards are working well for the institutions that use them, and are still needed.
- The Standards can be used in tandem with the Framework, instead of being discarded. There is room for both.
- Assessment is more challenging when employing a theoretical framework rather than specific standards.

Other IL practitioners also opposed revoking the Standards, citing that they are still of value in creating standardized, outcomes-based IL-assessment tests (Erlinger, 2018), as well as in one-shot, generic IL instruction (Gross et al., 2018). For the above reasons, many librarians in HE continue to base their IL instruction and assessment on the Standards (Erlinger, 2018; Sample, 2020), despite the valuable contribution of the Framework.

Another definition of IL worth mentioning was created by the UK's library and information association, CILIP. This definition is commonly used in Europe and describes IL across different sectors in society: "Information literacy is the ability to think critically and make balanced judgements about any information we find and use. It empowers us as citizens to reach and express informed views and to engage fully with society" (CILIP Information literacy group, 2018, p. 3).

### **Knowing, doing, and feeling**

Overall, being information literate involves skillful management of particular skills, knowledge, attitudes, values, behaviors, and aspirations. Put simply, it involves what I refer to in this dissertation as *knowing, doing, and feeling*. IL frameworks from around the world describe specific IL abilities, and the ACRL's IL Standards and Framework from the US have informed many of these. In Europe, knowledge, skills, and general competence are broadly described for all educational levels in each country's *national qualifications framework*<sup>3</sup>.

---

<sup>3</sup> A result of the Bologna process (1999), national qualifications frameworks ensure the compatibility of the quality and standards of HE across Europe, thus promoting cross-border mobility. European countries have compatible diplomas and learning outcomes for the three levels of HE: bachelor, master, and doctorate. Learning outcome descriptors in the national frameworks – categorized as knowledge, skills, and general competence – provide the basis for learning outcomes in individual subject curricula.



These frameworks, in turn, are incorporated into subject curricula. IL knowledge, skills, and competence, however, are less detailed here than in frameworks created specifically for IL, such as those from the ACRL.

IL *knowing* and *doing* are described in the ACRL Standard's outcomes and in the Framework's knowledge practices, using verbs such as know, identify, define, explore, describe, investigate, select, use, assess, organize, and demonstrate (ACRL, 2000, 2015). Most often, *knowing* is captured using tests, quizzes, or self-reports, while *doing* is captured with observations, exercises, or authentic assignments (Erlinger, 2018). Tests assess what students know, as opposed to what they do, and they can be used in either formative or summative assessment. *Feeling* encompasses the emotions, attitudes, and values associated with IL, and is described by the Framework's dispositions, and to a lesser extent in the Standards' outcomes, using verbs such as realize, acknowledge, consider, respect, value, and reflect on (ACRL, 2000, 2015). These aspects of IL are often captured with surveys or interviews (Erlinger, 2018). But how can we use these tools and methods more specifically to capture an individual's IL?

### **Measuring and assessing IL**

It is often recommended to design teaching based on the intended learning outcomes of the instruction (Løkse et al., 2017; Secker & Coonan, 2011). According to the principles of *constructive alignment* (Biggs, 2014), teaching and assessment methods are then aligned to these previously-defined student learning outcomes. If the alignment of learning outcomes, teaching, and assessment is successful, it should be possible to objectively measure whether an outcome, and thus student learning, is achieved. When designing IL instruction using the principles of constructive alignment, it follows that we must also have effective methods of measuring IL to assess student performance.

In the IL literature, the terms *measurement* and *assessment* are often intertwined. However, in the interest of clarity, I attempt to separate the two as much as possible in the review below, where *measurement* refers to tools and methods used to gauge an individual's performance, and *assessment* is a broader term that includes evaluating the individual on their achievement of a certain goal or objective (Calfee, 1993). Another source of ambiguity in the IL literature is that some assessment studies focus on assessing information-literacy instruction, while others focus on the students' attainment of IL learning goals. Since student learning can be indicative of the quality of the instruction, results for these two types of assessments – and the

literature regarding them – are often related. This will also be reflected in this literature review.

### *Measuring IL*

Students' IL knowledge and skills are commonly measured with either self-reports or tests (Boruff & Harrison, 2018). Self-reports in which students estimate their learning and competence may be subjective or misleading, as students – especially new students – tend to overestimate their abilities (Gross & Latham, 2012; Nierenberg & Fjeldbu, 2015; Nierenberg et al., 2021; Oakleaf et al., 2011; Schilling & Applegate, 2012). Accordingly, testing, with either locally developed or standardized IL tests, tends to measure knowledge and skills more accurately than self-report measures. IL-test items are normally multiple choice with predetermined correct responses and are based on curriculum learning outcomes or IL standards. Since creating a reliable and valid IL test is a demanding and difficult process, relatively few IL tests are psychometrically evaluated (Mahmood, 2017). Ideally, those who design IL tests, whether they are IL researchers or practitioners, should work to increase the reliability and validity of the tests.

Using articles published between 2007 and 2012, Schilling and Applegate (2012) compared the efficacy of commonly used measures for assessing IL instruction such as tests, surveys, graded assignments, and self-assessments. Their study discusses the pros and cons of various IL measures and finds disparities between the results obtained with these measures. The discrepancies can either be caused by shortcomings of the measures themselves, or by an actual disconnect between students' skills, attitudes, and behaviors, in other words, between what they *know*, *do*, and *feel*. One of Schilling and Applegate's conclusions is that IL learning is best assessed formatively, measuring changes in practical skills over time and providing feedback. They also posit that although affective measures – utilized for measuring attitudes, beliefs, and feelings – are commonly used in IL assessment, they are “not likely to provide meaningful evidence in terms of students' skills, course grades, or learning outcomes” (Schilling & Applegate, 2012, p. 266).

Mahmood (2017) performed a systematic review of evidence on psychometric properties of eighteen standardized tests for measuring IL knowledge and skills. Mahmood's review and other studies on IL testing (e.g., Hollis, 2018) demonstrate the need for more freely available, universal, validated, and reliable tools to measure IL. Mahmood reviewed studies from 2016

and earlier that reported the psychometric properties – including reliability and validity measures – of these tests. Of the eighteen tests, thirteen were developed in the United States and twelve were based on some or all five of the ACRL Standards, evidence of the leading role of the US in the field of IL (Mahmood, 2017). Of the four most widely used IL tests – iSkills (2006; discontinued in 2016), Research Readiness Self-Assessment (RRSA; 2004), Standardized Assessment of Information Literacy Skills (SAILS; 2001), and Information Literacy Test (ILT; 2007), only RRSA is freely available, while the others are commercial. Drawbacks with RRSA (Ivanitskaya et al., 2004), however, are its age and its limited focus on measuring IL in the field of health.

In the five years since Mahmood’s review, new psychometrically evaluated IL tests have been developed. However, independent reviews of their psychometric properties have yet to be conducted. One of these is the *Threshold Achievement Test for Information Literacy* (TATIL), a commercial test based on the ACRL Framework from 2015. Another is the Tromsø Information Literacy Test (TILT), developed by myself and my supervisors, which is freely available and based on IL standards (see Paper 1).

### *Assessing IL*

In literature from the field of education, assessment is often divided into two categories: *formative* and *summative*. Formative assessment takes place during instruction, providing feedback to learners that can be swiftly used to improve their learning. These assessments are typically informal and not graded. Summative assessments, on the other hand, are more formalized, and are often represented by grades. These can provide the learner with a comprehensive view of their learning after instruction is completed, providing them the opportunity to improve their future learning. Depending on the context, either formative or summative approaches can be used in IL assessment, and in some cases, it can be beneficial to use them in combination (Erlinger, 2018).

In addition to providing useful feedback and learning opportunities for students, IL assessments are also valuable to instructors and administrators. By demonstrating the efficacy of instructors’ methods, assessments can contribute to the development of their teaching. Furthermore, IL assessments can be used to demonstrate the value of IL programs for administrators and other stakeholders who grant the necessary funding.

Sobel and Sugimoto (2012) present various measures used to assess learning in IL instruction, how they are used by teaching librarians, and results of IL-assessment studies. Their exploratory study, based on a large survey of teaching librarians in academic libraries in the US, illustrates the range of teaching and assessment methods employed in IL instruction at these institutions. Sobel and Sugimoto found that a variety of assessment measures were in use, most of them formative. Measures included worksheets, pre- and post-session quizzes, self-assessments, tests, surveys, and graded assignments, and it was common for instructors to employ several different measures during the semester. Julien et al. (2018), in a similar survey of academic librarians in the US, found that student learning was assessed mainly through informal faculty feedback (58%), formative assessment during instruction (41%), and student self-assessments (40%). Note that the most prevalent form of feedback was from faculty, and it was informal. Julien et al. (2018) discovered that most IL instructors had not formulated specific learning goals for their instruction, thereby making formal assessment – and evaluation of their teaching – impossible.

The types of skills assessed by Sobel and Sugimoto (2012) were mainly information seeking, evaluation, and selection. Most of the IL tests documented in their study were informed by the ACRL Standards, however there was no consensus on how to operationalize outcomes from the Standards for assessment purposes, which would have allowed the comparison of assessment-study results (Sobel & Sugimoto, 2012). In the Julien et al. (2018) study, performed the year after the ACRL Framework was released, 41% of responding librarians indicated that the Framework had little or no influence on their IL teaching practice, while 31% reported that it had a major influence. In a systematic review of IL assessment studies conducted from 2006-2017 in the US and Canada, Erlinger (2018) found that the majority of published assessment studies and institutional case reports relied on the ACRL Standards for knowledge testing. At that point, she deemed it too early to detect the influence of the 2015 Framework in the IL-assessment literature.

Since Erlinger's systematic review, however, several Framework-based assessment studies have been published. One of these is Doyle et al. (2019), who designed a scale for measuring how students perceive their IL skills based on the ACRL Framework's knowledge practices and dispositions. Their scale measures seven distinct constructs of IL and shows evidence of reliability and validity. However, since it is a self-assessment, it is best suited for student learning when combined with skills-based assessment measures (Doyle et al., 2019).

Another study (Clarke & Radcliff, 2018) compares locally developed IL tests to standardized, commercial IL tests, all grounded in the ACRL Framework. According to Clarke and Radcliff (2018), there are advantages and disadvantages of both types of tests. Locally developed tests align with the institution's objectives and outcomes and are free of cost, however, they are rarely psychometrically evaluated. Standardized tests, on the other hand, are evaluated for reliability and validity, but can be expensive. In addition to reflecting national standards and practices and allowing comparisons between institutions, standardized tests may also align to some extent with local learning outcomes (Clarke & Radcliff, 2018).

Erlinger (2018) uncovered some general challenges with IL assessment in her review. First – and contrary to Schilling and Applegate's (2012) observations – Erlinger found that although cognitive and behavioral aspects of IL are frequently studied, the *affective* domain – including emotions and motivations – is much less emphasized despite its importance to the development of IL (Erlinger, 2018). Second, although it is important that assessment measures are used to detect changes in IL abilities over time, for example with pre- and post-tests (Sobel & Sugimoto, 2012), very few longitudinal studies have been performed. Most pre-/post-studies in the IL literature measure only short-term learning over a semester or school year, rather than longer-term learning over several years (Erlinger, 2018). Third, few IL assessment measures are psychometrically evaluated (Erlinger, 2018; Mahmood, 2017), although this is essential in order for their results to be considered valid and reliable (American Educational Research Association, 1999).

### **Dimensionality of the IL construct**

Depending on the definition, standard, framework, or model consulted, IL can encompass many different facets. These include, for example, recognizing an information need; seeking, locating, accessing, critically evaluating, organizing, citing, and communicating information; using information to create new knowledge; and understanding cultural, ethical, economic, legal, and social issues surrounding the use of information. In addition – making the construct even more ambiguous – IL can be described by, or integrated into, several different terms, each with its own perspective. These include media literacy, digital literacy, information skills, computer and information literacy, and metaliteracy.

Despite this multitude of descriptions and labels, IL is regarded as a unidimensional construct in much of the literature (e.g., Hollis, 2018). Those who work with IL seem to understand intuitively what it is and to consider it a distinct field. However, among researchers who are

concerned with the dimensionality of the IL construct, there is a lack of consensus regarding its homogeneity. Several IL scholars, including Beile O'Neil (2005) and Morley (2014), failed to find distinct facets in their tests when investigating the dimensionality of the IL construct with factor analyses. This debate has implications also for the assessment of IL. Is it appropriate to create a standard measure for IL in its entirety, or is it more reasonable to assess its facets/elements individually?

Despite the lack of consensus regarding IL's definition, dimensionality, and assessment, IL as a field of practice and a domain of research has grown substantially since its inception nearly 50 years ago. As well as being the subject of research and practice in educational settings, IL is the focus of several conferences and journals around the world. With the deluge of mis- and disinformation in recent years – regarding for example COVID-19 vaccines, the war in Ukraine, and the 2020 US election results – the importance of being information literate has become increasingly apparent for everyone, also beyond the educational context.

### **1.2.2 Transformative learning theory**

As previously mentioned, the ACRL Framework's core concepts were informed by ideas of *threshold concepts* (ACRL, 2015). These threshold concepts were in turn informed by various learning theories, including the theory of *transformative learning* (Meyer et al., 2010), originally introduced by Jack Mezirow.

In this section of the literature review, I introduce some of the scholarly literature pertaining to transformative learning (TL) and its relevance for IL learning. Learning theorists such as Jack Mezirow, Paulo Freire, Robert Kegan, and Knud Illeris have slightly different perspectives of TL. Other scholars, such as Baartman and de Bruijn, have learning theories that can be considered theories of TL, although they do not use the term *transformative learning*. I concentrate in this section on the theories of Mezirow, Illeris, and Baartman and de Bruijn.

Jack Mezirow (1923-2014) is considered the founder of TL theory. Mezirow (1978, 1994, 2000, 2003) developed this theory to understand and describe the learning process and transformation that can occur when adults re-enter higher education after experience elsewhere. Mezirow's interest in this topic was sparked in 1978 when he observed his wife's personal transformation after returning to college in midlife. This transformation was typical

for many American women during women's liberation movement in the 1970s, as women became more conscious of, and liberated themselves from, societal and personal oppression.

Mezirow was also inspired by Paulo Freire's (1970) related work, grounded in critical pedagogy, exploring the oppression, consciousness-raising, and education of illiterate agricultural workers in Brazil. Both Freire and Mezirow were experiential education theorists, a teaching philosophy that stems from critical theory (Downey, 2014) and emphasizes learners' self-reflection and life experience in their education. The assumptions of TL are thereby constructivist (Mezirow, 1994), a theory of education that recognizes that individuals *construct* meaning and learn by interpreting their life experiences, thereby creating new knowledge and understandings.

Mezirow's TL theory posits that our past experiences represent significant learning opportunities, and that critical self-reflection and questioning of these experiences can lead to a transformation of our perspectives and understanding (Mezirow, 1978, 1994, 2000, 2003). When presented with new information, we evaluate and reinterpret our past assumptions and understandings, examining ideas with different perspectives to create new knowledge, insight, and meaning.

TL is described by Mezirow as a theory of adult learning (andragogy), especially relevant to adults returning to education later in life. However, since Mezirow (2000, pp. 24-25) defines an adult as "a person old enough to be held responsible for his or her acts," many researchers in HE study TL also in college students (Christie et al., 2015; Hooper & Scharf, 2017). According to the main principles of andragogy, adult learners are more self-directed, they base their learning more on previous experience, and they apply their learning more to practical problems than younger learners (Knowles, 1988). Since Mezirow first introduced TL, college students have become increasingly responsible for their own learning and education has become more problem-based. For the above reasons, the theoretical difference between adult education and higher education in general has been debated in recent years (Wang, 2017), and the two have become less clearly separated in the TL literature (Christie et al., 2015).

The development of TL theory began in the late 1970's, when Mezirow proposed the concept of *perspective transformation*: "a structural change in the way we see ourselves and our relationships" (Mezirow, 1978, p. 100). This became an important component of his TL

theory, introduced a decade later. At that time, Mezirow defined TL as “the process of learning through critical self-reflection, which results in the reformulation of a *meaning perspective* [one's assumptions when interpreting the meaning of an experience] to allow a more inclusive, discriminating, and *integrative* understanding of one's experience” (Mezirow, 1990, pp. xvi, my italics and brackets). He added that TL can include taking actions based on these new insights. Mezirow further developed his definition of TL at several stages. In 2003, he added the concept of *frames of reference*, which encompasses our values, language, culture, and personal preferences, and includes an emotional component (Mezirow, 2003). He wrote that “transformative learning is learning that transforms problematic frames of reference – sets of fixed assumptions and expectations (habits of mind, meaning perspectives, mindsets) – to make them more inclusive, discriminating, open, reflective, and emotionally able to change” (Mezirow, 2003, p. 58). Mezirow maintained that these frames of reference, due to their increased ability to generate new opinions and beliefs, provide a better basis for guiding action than the previous meaning perspectives. The *habits of mind* he refers to are the broad assumptions from which we view our surroundings.

An important part of Mezirow's (1990) original definition of TL is that it involves the *integrative* understanding of an individual's experience. The process of integration implicates building relationships between knowledge, skills, and attitudes (Baartman & de Bruijn, 2011), which I study in relation to students' IL in this dissertation. For example, when writing a term paper, a student may integrate *knowledge* (about how to search for relevant literature, or how to cite sources when writing), *skills* (identifying and selecting relevant literature, using sources to develop and support their own argument), and *attitudes* (wanting to become information literate, recognizing the limits of their IL skills, following legal and ethical guidelines when using information). Baartman and de Bruijn (2011) describe a type of integration that they call *transformative integration*, which can be interpreted as TL that is brought about by the process of integration. They conceive of transformative integration as involving critical reflection of the knowledge, skills, and attitudes necessary to perform a task. This requires using time to reassess one's perspectives and to transform them if the new information does not fit with one's previous mental models or conceptions (Baartman & de Bruijn, 2011). This implies that transformative integration is a type of TL. Transformative integration involves a learning process called *accommodation* (Baartman & de Bruijn, 2011). Accommodation occurs when “one breaks down (parts of) an existing scheme or pattern and *transforms* it so that the new situation can be linked in” (Illeris, 2004, p. 84, my italics). By



using the word *transform* in his definition of accommodation, Illeris implies that accommodation may also be a TL process in some cases. An example of accommodation would be a student who previously thought that highlighting text was the best way to learn new material, but recently learned in a psychology lecture that highlighting is actually much less effective than other learning strategies. She would then have to accommodate her perspective on learning and perhaps change the way she studies. In more extreme instances, for example if a student gets expelled from the university for plagiarizing, a different process of learning might take place, requiring more than accommodation. When this learning involves critical self-reflection and the restructuring of cognitive, emotional, and social factors, Illeris (2004) considers it a transformative learning process.

Although originally introduced over 40 years ago, TL is still considered a dominant theory of adult learning internationally. Several scholars however, including Illeris (2014a, 2014b, 2017) and Kegan (2000), criticize TL theory for being too cognitively focused. They argue that Mezirow's definition of TL is not sufficiently distinguishable from information processing, a mainly cognitive learning process. Other scholars, including Hoggan (2016), criticize Mezirow's definition and theory of TL for being too broad and lacking clarity. The term can be used to refer to nearly all adult learning outcomes instead of just the profound transformations that significantly change people's lives, thus reducing the theory's usefulness and value (Hoggan, 2016, p. 60). However, despite general agreement in the field of andragogy that the original definition of TL was insufficient (Illeris, 2017), no alternative definition was suggested for many years.

### **1.2.3 Identity**

In 2014, however, Illeris proposed a new definition of TL that includes the keyword *identity*. After evaluating various terms, Illeris (2014b) concluded that *identity* was the best to describe what it is that TL actually transforms, and to capture the multitude of dimensions – including psychological and sociological – that comprise learning as a whole. In Illeris's definition, TL comprises “all learning which implies changes in the identity of the learner” (Illeris, 2014b, p. 40). Illeris also claimed that since the word *identity* is used in everyday language, it is immediately understood by everyone (although some may dispute how accurate people's understandings of the term actually are).

Illeris (2017) maintains that the inclusion of identity in TL theory promotes new, enriched understandings of both concepts, and provides greater insights into learning, transformation,

and empowerment. Illeris (2017) believes that previous definitions of TL lack important *emotional* and *social* dimensions necessary to understand and explain personal transformations. In addition, Illeris (2014a, p. 36) points out that Mezirow does not fully account for the situatedness of learning in his TL theory, although he certainly recognizes it to some extent. Illeris (2017) posits that identity's inclusion in TL improves Mezirow's theory in two ways: (1) by adding much-needed psychosocial elements to TL's mainly cognitive orientation, and (2) by accounting more for the situatedness of learning in society. Illeris (2014a, p. 36) maintains that – although Mezirow saw the importance of emotional and social elements – he did not consider them to be something (in addition to cognition) that gets transformed in TL, and instead defended the primacy of cognitive elements. However, according to Illeris (2017), these psychosocial elements are of value in describing and analyzing the perspective transformation that leads to personal development, including changes in attitudes and behavior.

Erik Erikson, an influential psychoanalyst known for his extensive theoretical work on identity, and who coined the term *identity crisis*, must be mentioned in any discussion of identity. He describes personal identity as being composed of two elements: how we perceive ourselves as being the same in all circumstances, and our perceptions of how others judge us (Erikson, 1968). Erikson maintains that personal identity is formed during youth – early teens to late twenties – and remains relatively stable throughout life.

Gradually however, sociologists such as Zygmunt Bauman and Anthony Giddens began to question Erikson's belief in the lifelong stability of personal identity. Bauman (2000) describes the world today as being in a state of *liquid modernity* which leads to continual, fluid changes in identities and relationships, while Giddens (1991) posits that our identities change continuously in order to sustain our sense of self. These modern understandings of identity as flexible rather than fixed thereby support Illeris's supposition that identity is something that can be transformed in TL.

Illeris (2014a, chap. 9) created a model for the internal structure of identity, in which its three main elements are depicted as layers. At the center is core identity, described as “the experience of being a distinct and particular individual, being the same in different situations and connections, and also through the entire life course, in spite of any development or change” (Illeris, 2014a, p. 70). Illeris's *core identity* is thereby similar to Erikson's description of *personal identity*. Changes to core identity can occur, however, either gradually

or suddenly, for example when identifying as a parent after having a child (Illeris, 2014a, p. 72).

The middle layer in Illeris's model – the personality layer – is more flexible than the core identity layer and can change in response to new conditions or impressions. This layer includes typical personality traits such as empathy, behavior, attitudes, and values, and describes “how one relates to others, communities and groups, important issues and instances, significant events, and incidents – overall to the outside world, society, and environment of which one is a part” (Illeris, 2014a, p. 72). Illeris points out that it is within the personality layer that Mezirow's meaning perspectives and frames of reference are situated, and that changes here are the common target of TL.

The outermost layer in the model is the preference layer, the least stable of the three identity elements (Illeris, 2014a, p. 73). In this layer are our instinctive preferences, routines, and how we think, feel, and act in situations we are less engaged with. Since identity as a whole is not significantly affected by changes in the preference layer, TL is not normally relevant here.

#### *Relevance of transformative learning and identity to IL*

Importantly for this dissertation, the theoretical framework of TL is applicable to learning in many fields, including information literacy. TL was first introduced to the field of LIS in 2008 in a dissertation entitled *The Transformative Library* (Kenney, 2008), where TL informed Kenney's model for information use among public library users. Many IL scholars, including Hooper and Scharf (2017), Nichols Hess (2018), Whitworth (2012), and Kos and Špiranec (2014), maintain that TL is also relevant to the IL instruction offered by academic libraries and to theories of LIS in general.

According to Mezirow (1990, 1991), both critical thinking and critical reflection are necessary components of TL. Although they can be viewed as quite similar, critical thinking involves seeking reasons on which to base one's opinions, beliefs, assessments, and actions, while critical reflection involves questioning the validity of one's values and beliefs (Bates, 2019; Mezirow, 1990). Being information literate involves both, for example when evaluating the quality and reliability of individual information sources (critical thinking), or when questioning the power structures involved in disseminating those information sources, which

may be oppressive (critical reflection).<sup>4</sup> Critical thinking and critical reflection are both important for lifelong learning (Hooper & Scharf, 2017). Critical reflection may be a mainly outward process where the learner critiques the assumptions that underlie an issue, or a more inward process of *self*-reflection regarding the learner's values and beliefs (Nichols Hess, 2018).

It is obvious from the introduction to the ACRL Framework that its threshold concepts were inspired by TL theory. These threshold concepts are described as:

... core or foundational concepts that, once grasped by the learner, create new perspectives and ways of understanding a discipline or challenging knowledge domain. Such concepts produce transformation within the learner; without them, the learner does not acquire expertise in that field of knowledge. Threshold concepts can be thought of as portals through which the learner must pass in order to develop new perspectives and wider understanding. (ACRL, 2015, p. 9)

As seen in both the language and content of this definition, these threshold concepts are grounded in TL theory's perspective transformations and are constructivist in nature. Although there is no direct reference to Mezirow, the Framework cites Meyer et al. (2010) who refer to Mezirow in their discussion of threshold concepts.

According to Kenney (2008), TL's shift between an "unexamined way of thinking to a more examined and critically reflective way," (Mezirow, 2003, p. 59) is an internal and largely invisible process, and thereby difficult to measure with testing, rubrics, or in outcomes-based education in general. With regard to IL learning, it therefore follows that skill-based outcomes such as those provided by the ACRL Standards may be insufficient in detecting the types of personal transformations described in TL (Kos & Špiranec, 2014). These outcomes focus on cognitive, rather than psychosocial, aspects of learning, as is reflected by the Standards' previously mentioned verbs *know*, *identify*, *define*, and *organize*. Tests and questionnaires based on these outcomes have thereby limited ability to detect TL's perspective

---

<sup>4</sup> The development of a critical consciousness about information is known as *critical information literacy* (CIL). The ACRL Framework (2015) draws on the premises of CIL, and many academic librarians use CIL's theoretical framework to inform their IL instruction.

transformations. The ACRL Framework's threshold concepts, knowledge practices, and dispositions, on the other hand, using verbs including *realize*, *acknowledge*, *value*, and *reflect on*, may be better equipped to assess TL (Hooper & Scharf, 2017).

Several scholars have examined TL theory and identity transformations in the field of IL, including Whitworth (2009, 2012), Kos and Špiranec (2014), and Nichols Hess (2018, 2019, 2020). In their research, they mainly explored the perspective transformations and identities of IL practitioners (academic librarians), including how they are perceived by others, rather than investigating transformations in students. Whitworth (2012) maintains that reflective practice among IL practitioners is necessary to achieve transformative IL education, in which librarians are perceived as teachers in the effective *use* of information rather than solely as *providers* of information. To develop this reflective practice, IL practitioners need to continuously reflect on the relationship between theory and practice in their teaching (Whitworth, 2012). One example of this could be recognizing the value of integrating discipline-specific information skills in their instruction. Another example could be encouraging learners to critically reflect both on their information sources *and* on their information searches.

Kos and Špiranec (2014) also recognize a connection between IL and transformative learning. In order to achieve TL's critical thinking and critical reflection, people must first acknowledge their underlying assumptions. This in turn requires the ability to find and access reliable and relevant information – a major tenet of IL – whose content may be at odds with one's previous beliefs. Nichols Hess (2018, 2019, 2020) explored whether academic librarians who teach IL develop teaching identities through transformative experiences, and she found that some did. To measure their perspective transformations, Nichols Hess used a revised version of the *Learning Activities Survey*, a popular survey for studying TL in educators developed by K. P. King (2009).

In my doctoral project, rather than studying transformations in those who *teach* IL, I attempt to detect transformations in those who *learn* IL.

#### **1.2.4 Interest theory**

Interest fuels learning and growth with any object of interest (Akkerman & Bakker, 2019), including IL. The association between IL and interest has been very little explored previously, although student interest can have significance for how we teach IL.

Interest is an important factor in learning theories, and its role in learning has been extensively studied by educational psychologists including T. I. Dahl, S. Hidi, A. Krapp, K. A. Renninger, J. I. Rotgans, G. Shraw, and P. J. Silvia. Based on various interest models and corresponding theoretical perspectives – from the above-mentioned scholars and others – several methods have been developed for measuring interest, varying in both content and form. Some are based on observation, while others rely on self-reports with a variety of theoretically derived items, subscales, or scoring schemes (for an overview, see Paper 2). While some measures are specifically designed for one particular object or domain of interest, such as mathematics, others are tailorable to different domains. The great variation in interest measures makes it difficult to compare research results, and a gold standard for measuring interest is therefore desirable.

One persuasive, dynamic model of interest that could potentially provide the theoretical foundation for such a standard, is Hidi and Renninger's (2006) *four-phase model of interest development*. Built upon empirical studies of learning and interest, this model recognizes that interest is something that can wax or wane over time; just because you are currently interested in baking sour-dough bread, does not mean that you will be next month. According to this model, some interests are fleeting, occupying only a slice in time, while other interests are more stable over time. This has important implications for learning, as people with stable, personally sustained interests re-engage more readily with their content and are therefore more likely to learn that content better (Hidi & Renninger, 2006).

Effective learning is accomplished not only by accumulating knowledge and skills (cognition), but also by creating positive emotion, interest, and personal relevance, and thus *meaningfulness*, in relation to the content (Hidi & Renninger, 2006). In the context of HE, Hidi and Renninger (2006) define interest as a student's predisposition to re-engage in a subject or activity and in terms of the corresponding psychological state during this process.

To capture interests' stability and fluctuations, Hidi and Renninger (2006) systematized interest into four sequential, qualitatively distinct phases. These phases represent different psychological states, with varying degrees of situational dependence, positive affect, competence, and meaningfulness. In the first two interest phases, interest is situational, supported by environmental stimuli in the moment. In the last two phases, interest is individual, and more personally sustained. Having interest in something, whether it is

situational or individual, has been proven to positively affect cognitive performance and to improve learning (Hidi & Renninger, 2006, p. 13).

More specifically, the four phases of interest are:

1. *Triggered situational interest* – interest is triggered by a particular situation or sparked by something that catches attention. It is fleeting and has short-term positive effects on emotion and cognition. Students in this phase need feedback and encouragement to maintain their interest, and can be easily overwhelmed by too much information.
2. *Maintained situational interest* – focused attention persists over a longer period but is still dependent on the situation. It feels more meaningful to the learner than triggered situational interest and it generates more positive feelings. Students begin to explore ideas on their own in this phase, but still need external support and encouragement. They may want direct instruction and concrete suggestions about what to undertake.
3. *Emerging individual interest* – the person is driven to re-engage with this interest over time, independent of the situation. It is associated with positive affect, and it is typically self-generated. It requires some external support to be sustained, but less than in phases 1 and 2.
4. *Well-developed individual interest* – the person has a persistent tendency to re-engage with this interest over time, since it has more value and creates greater positive affect than interests in phases 1-3. It is more long term and self-generated, and the person will pursue this interest despite challenges. The person is more reflective about the content and appreciates feedback.

Hidi and Renninger, however, did not operationalize their four-phase model to enable testing and validation. Empirical studies attempting to differentiate between the model's interest phases have not had sufficient reliability or theoretically based construct validity to do so (Dahl & Nierenberg, 2021). Until now, with our own study (Paper 2), Hidi and Renninger's four-phase model of interest development had not been fully operationalized and tested.

### **Interest and IL**

With the exception of Black (2018), little has been written about the connection between IL and interest. Black refers to Renninger's (2009) model of interest and identity and to Hidi and Renninger's (2006) four-phase model of interest development, and discusses the models' relevance to IL instruction. In the context of HE, Renninger (2009) defines identity as a

student's self-image as someone who engages in certain activities or content. This definition falls within what Illeris would call the personality layer of identity, where changes occur in response to new impressions or conditions. Identity is linked to interest in that people's self-images are often related to their interests and to corresponding phases of interest development (Renninger, 2009). For example, the degree to which one identifies oneself as a skier, a bird-watcher, or an information-literate person, depends on how developed one's interest is in these activities/conditions.

Depending on their phase of interest, learners need and want different kinds of feedback and challenges from their teachers (Hidi & Renninger, 2006; Renninger, 2009). This has consequences for how we teach, and underscores the importance of interest for learning and identity development. For example, a newly triggered interest in IL can be sustained by presenting content that is meaningful for learners, for example by showing them how to find and access peer-reviewed literature for their upcoming assignment. Librarians who teach IL could benefit from being familiar with theories of interest – such as Hidi and Renninger's (2006) four-phase model of interest development and Renninger's (2009) model of interest and identity – when designing IL instruction, in order to best arouse, maintain, and sustain the development of student interest in IL (Black, 2018).

### **1.2.5 Gaps in the literature**

As we have seen in this review, there are several gaps in the IL literature that warrant further investigation. The following points enumerate research needed to understand the development of IL in students, and which represent the major themes of this doctoral project.

- a) There is a demand for freely available, psychometrically tested, and easily implemented tools to measure IL knowledge, skills, and attitudes. These measures should be as universal as possible, and thereby reliable and valid for most disciplines and geographic locations. Few existing IL tests are psychometrically evaluated (Mahmood, 2017), yet this is essential in order for both research and learning assessment results to be considered valid and reliable, and to use measures for comparison purposes.
- b) With the possible exceptions of Schilling and Applegate (2007, 2012) and Pinto and Fernández-Pascual (2019), few studies measure IL *knowing*, *doing*, and *feeling*, and the associations among them, in students. Most studies focus on one or two of these, yet all are contributing factors to IL learning. Feelings, including emotions and



interest, are less emphasized in the IL literature despite their demonstrated importance for motivation and learning.

- c) Although several studies have been published that employ tests to measure IL before and after interventions, longer-term studies are sparse in the scholarly literature. There is a clear need for studies that measure IL development over time, as these are more effective than short-term studies for gauging long-term retention and evolution of IL knowledge, skills, and attitudes.
- d) The connection between IL and interest has barely been explored in the IL literature. Studies investigating the association between interest and IL can help us to understand how interest impacts motivation and learning, and could be used to inform how we teach IL to students with different degrees of interest in IL. Such exploration involves examining several factors, including phases of interest development as well as the associations among students' interest, their felt need to learn more IL skills, and the effort they intend to put into that learning.
- e) The degree to which transformative IL learning takes place in undergraduates – as measured by the degree to which students' identities change as a result of becoming information literate – needs further investigation. Although not explicitly stated in the ACRL Framework, its threshold concepts imply that such TL is a desired outcome. Previous research on transformative learning in IL focuses on the transformative experiences of IL practitioners rather than students.
- f) The degree to which students are metacognitively aware of their IL abilities has been studied to some extent, but further investigation is necessary to determine precisely which factors – such as gender, education level, performance level, and timing (pre- vs. post-test estimates) – influence their metacognition in ways that may matter to IL development.

### **1.3 Research questions**

The overarching research question addressed in this dissertation is: *How do students develop as information-literate individuals in the first three years of their undergraduate education?*

Based on this main research question and the aforementioned gaps in the literature, three studies were performed, each with its own subset of research questions. These studies, and their coherence to the main research question, are briefly described below.

- Study 1 (Papers 1 and 2): How can we best develop valid and reliable tools to effectively measure IL knowledge (*know*), skills (*do*), and attitudes (*feel*)? Tools for measuring IL knowledge and skills are addressed in Paper 1, while tools for quantitatively measuring IL attitudes – defined in this research as students’ interest in being or becoming information-literate people – are addressed in Paper 2. In addition, Paper 1 addresses the dimensionality of the IL construct. Is IL a unidimensional construct? If not, what implications does this have?

Coherence: In order to determine how students develop as information-literate people, we need valid and reliable tools to measure various aspects of their IL.

- Study 2 (Paper 3): How self-aware are students of their own IL levels before and after taking an objective IL-knowledge test? Which other variables may influence metacognitive accuracy? What is the relationship between students’ interest in being or becoming information-literate people, their felt need to learn more IL skills, and their intent to do so?

Coherence: Metacognition involves self-reflection regarding one’s cognitive abilities. Metacognitive assessment is an important part of learning, and this study thereby provides a valuable addition to Study 1’s objective assessments of IL learning in this project.

- Study 3 (Paper 4): In which ways do students develop as IL individuals over the first three years as undergraduates? How do relationships evolve among what students *know* about IL, what they *do* in practice in authentic assignments involving IL skills, and what they *feel*, specifically their interest in being or becoming information-literate people? How do students’ regard their own development as information literate individuals over the three years? Do their identities change as a result of becoming more information literate? Does TL take place?

Coherence: The tools developed in the first study are used to assess IL development in the third study, and interviews were conducted to examine students’ own perceptions of their IL development. Both objective and subjective results of the three-year study of undergraduates’ IL development are presented here, directly addressing the main research question.

## 2 Methods

This doctoral dissertation is composed of four papers, as well as this extended summary which ties the papers together as a coherent whole. The first three papers are quantitative, while the fourth combines and integrates quantitative and qualitative methods.

Methodologically, the dissertation as a whole therefore employs mixed methods. Details of the methods are described in the four individual papers, while general and overarching methodological questions and short summaries of the methods used are addressed in this extended summary.

Several individual studies were conducted in this doctoral research. When referring to the entirety of the research, I therefore use the term *project* instead of *study*.

In the preceding literature review, I positioned my research to the current state of knowledge about IL-relevant concepts. In the Methods sections, I position my research ethically, theoretically, methodologically, and in relation to the philosophy of science.

### 2.1 Research ethics

Research in the social sciences involves collecting data. It is important that data are collected, processed, stored, and managed in an ethically responsible manner. There are strict ethical guidelines and procedures for research in Norway, intended to protect the privacy of participants, to promote good scientific practice, and prevent scientific misconduct. I have adhered to the following guidelines and procedures in my research:

- [General guidelines](#) from The Norwegian National Research Ethics Committees
- [Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology](#) from The National Committee for Research Ethics in the Social Sciences and the Humanities
- [Guidelines by The Norwegian National Research Ethics Committee for medical and health research](#) (See Appendix 2 for course certificate, in Norwegian.)
- [Notification form for personal data](#) from The Norwegian Centre for Research Data (NSD). (See Appendix 3 for notification form sent to participants, in Norwegian.)
- Consent form for participants (See Appendix 4 for the consent form for the student group that was followed longitudinally. In Norwegian, assessed by NSD.)

- [Data management plan](#) from NSD. (See Appendix 5 for the plan for my study. Translated from Norwegian. Note: changes have been made since the outset of the project.)

## 2.2 Theoretical stance

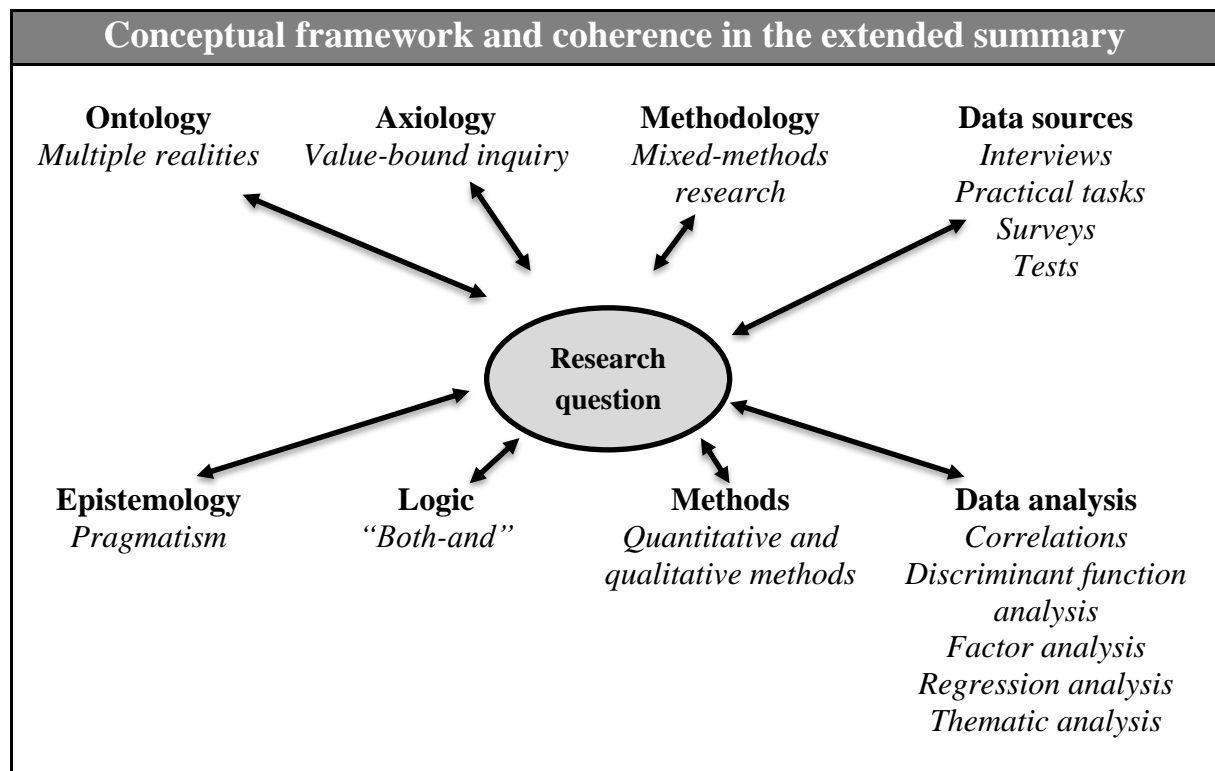
The theoretical stance taken by researchers is informed by their philosophical assumptions about knowledge, reality, and values. Epistemology can be described by asking: *What is knowledge? How do we know what we know?* Answers to these questions underpin what we consider to be valid knowledge in our field, and can be placed on a continuum from objective to subjective points of view (Pearce, 2015, p. 44). Ontology relates to the question: *What is the nature of reality?* The ontological spectrum ranges from the belief in an objective reality determined by cause and effect (positivist or postpositivist), to the belief in several constructed realities determined by meaning and interpretation (interpretivist; Biesta, 2010, p. 9). Axiology asks: *What is the role of values in research?* Axiological positions range from value-free (positivist) to value-bound (constructivist).

In this IL project, the underpinning epistemological assumptions and research paradigm are pragmatic (see Figure 1). In a pragmatic perspective, several forms of logical inference – for example, inductive, deductive, and abductive – are used to generate the most likely conclusion or best explanation based on various sources or methods. Tashakkori et al. (2020) call this type of logic, *both-and reasoning*.

Pragmatism involves grounding research in a specific situation, and frequently combines several ways of describing reality (Pearce, 2015, p. 44). Ontologically, this approach is underpinned by the assumption that multiple kinds of reality exist; reality is neither purely objective, purely intersubjective, nor purely subjective (Tashakkori et al., 2020, p. 71). Different disciplines can identify and describe, for example, different social, psychological, and economic aspects of reality (Tashakkori et al., 2020, p. 64). This is often accomplished by a combination of quantitative and qualitative methods (McEvoy & Richards, 2006), for example with a mixed-methods methodology.

Axiologically, values in a pragmatic paradigm are guided by what we study and how we interpret our results with the goal of answering our research questions (Tashakkori et al., 2020, p. 63).

Figure 1: Pragmatism Paradigm: Conceptual Framework



### 2.3 Methodology – Mixed-methods research

The choice of methodology should always be driven by the research question. At times, the goal of research may be to test and confirm hypotheses, while at other times, the goal may be to explore and discover. Testing and confirming of hypotheses is normally accomplished using quantitative (QUAN) methods, while exploring and discovering is commonly the realm of qualitative (QUAL) methods. These distinctions, however, are not absolute; there is no firm line between quantitatively driven and qualitatively driven approaches, but rather, they are on a continuum (Hesse-Biber et al., 2015, p. 4). QUAN analyses are suitable for demonstrating the strength of associations, while QUAL results show the nature of those associations (Fetters et al., 2013).

QUAN research dominated the social and behavioral sciences in the twentieth century (Tashakkori et al., 2020). It relies primarily on numerical data and analyses, typically to test hypotheses. These hypotheses are generally based on theory or previous research. Using statistical analyses (based on probabilities rather than absolutes) and deductive reasoning, QUAN researchers determine whether their predicted hypotheses can arguably be said to be supported by the data. QUAN research is generally associated with a positivist or

postpositivist paradigm. In a positivist paradigm, research is considered objective and uninfluenced by the values of the researcher, whereas in a postpositivist paradigm, researchers acknowledge that their values can affect their research. Most QUAN research in the social and behavioral sciences today has a postpositivistic orientation (Tashakkori et al., 2020, p. 5).

QUAL research, on the other hand, is often associated with a constructivist paradigm (Tashakkori et al., 2020, p. 7). In a QUAL paradigm, researchers commonly use more inductive reasoning to construct theory based on data, for example, narrative data from interviews or open-ended survey questions. In this case, thematic analysis can be used to analyze the data; the first stage typically produces themes and the second stage attempts to find complex relationships in the data (Tashakkori et al., 2020, p. 7). The units of analysis in QUAL research, however, can be based on quite varied sources – not just narrative data from interviews – and there are multiple methods available to analyze them.

There are advantages and disadvantages to both QUAN and QUAL research paradigms. QUAN study designs, though specific in their purposes (e.g., testing and confirming) and exact in their means of measurement, may, for those same reasons, limit our ability to make new discoveries or explain phenomena in novel ways. QUAL methods, on the other hand, while allowing us to explore and discover, may limit our ability to generalize our findings to other settings (Pearce, 2015, p. 46).

A third research paradigm that has emerged in recent years – *mixed methods* – involves deliberately combining QUAN and QUAL approaches. Johnson et al. (2007, p. 123) define mixed methods as research that:

... combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration.

This *methodological pluralism* enhances understanding of complex phenomena (Johnson & Onwuegbuzie, 2004), and mixed methods has therefore become popular in complex disciplines such as the social and behavioral sciences (Tashakkori & Teddlie, 2016).

In mixed-methods research, where multiple methods are employed in the quest for knowledge, the strengths in one method can compensate for weaknesses in the other. This approach therefore allows us to investigate complex topics more comprehensively by

approaching them from several angles. While mixed methods has become increasingly prevalent in LIS studies (Granikov et al., 2020), our search in Library & Information Science Abstracts (LISA) revealed few mixed-methods publications focusing on IL.

Mixed-methods research supports a pragmatic approach, allowing a mixture of QUAN and QUAL methods in data collection and analyses to find the best explanation, or most likely conclusion, to a research question (Onwuegbuzie et al., 2009). A mixed-methods methodology therefore aligns with this project's theoretical stance and underpinning epistemological, ontological, and axiological assumptions.

The tenets of mixed-methods research are continually evolving. In this dissertation I consult and refer mainly to one of the latest methodological works in mixed methods, namely the second edition of *Foundations of Mixed-Methods Research*, by Tashakkori, Burke Johnson, and Teddlie (2020).

One important feature of mixed-methods research is that it is based on the principle of *integration*; if the approaches employed in a study are not integrated, it should not be considered mixed methods (Tashakkori et al., 2020, p. 120)<sup>5</sup>. The integration of QUAN or QUAL methods enhances the credibility of findings, as results can clarify, corroborate, and complement each other (Schoonenboom & Johnson, 2017).

## **2.4 Research design – Multisample Hybrid Mixed-Methods Design**

A study's research design must align with both the research questions and the chosen methodology, which in this project is mixed methods. Mixed-methods research designs require flexibility and creativity from researchers who strive to combine QUAN and QUAL methods in a way that most effectively answers their research questions (Tashakkori et al., 2020, p. 114). Each phase or component in a mixed-methods study is called a *strand*. In mixed-methods studies, it is common for the research design to evolve with time as researchers collect and analyze data. Some research questions are *preplanned*, while others

---

<sup>5</sup> Tashakkori et al. (2020, p. 120) suggest that study designs in which QUAN and QUAL data are collected and analyzed separately, without integration, be considered “quasi-mixed methods.”

are *emergent*, created during the course of the study as needed (Tashakkori et al., 2020). The final research design may therefore differ from the original plan – as was the case in this project – with emerging QUAL or QUAN components added to the preplanned components as found necessary in order to answer the research questions most effectively.

Although this IL project originally relied on QUAN methods, the possibility of employing QUAL methods in addition was never excluded. When it became apparent that QUAN methods alone were limited in their capacity to convey the full story of students' IL development, a new research plan gradually emerged. To the QUAN measures – a multiple-choice test, a closed-answer questionnaire, and two assignment-based tasks – I added QUAL measures, namely interviews. We performed both individual and focus-group interviews to provide students the opportunity to reflect on their own learning and aspects of their IL development. The purpose of the interviews was to explain and enhance the QUAN findings, adding depth to our understanding of their IL development.

According to Fetters et al. (2013), there are three basic study designs in mixed-methods research: exploratory sequential designs, explanatory sequential designs, and convergent designs. In *exploratory sequential designs*, research begins by collecting and analyzing QUAL data, and the results inform subsequent QUAN data collection and analysis. Studies with *explanatory sequential designs* are performed in the opposite order, with findings from the QUAN data informing subsequent QUAL data collection and analysis. With *convergent designs*, both types of data are collected and analyzed simultaneously (Fetters et al., 2013).

The mixed-methods research design in this project is what Fetters et al. (2013) might label *explanatory sequential design*, where the main QUAN component informs and is succeeded by a supplemental QUAL component. However, given the chronology and complexity of data collection in this IL project (see Table 1), these three labels are misleading in their simplicity. QUAN data, from several different samples, were collected both before, between, and after the QUAL data collection (grey rows in Table 1). Additionally, because of the relatively large number of strands in this long-term study, these labels do not fully reflect the influence of data, results, and inferences from previous strands on subsequent survey and interview design.



Table 1: Data Collection, Listed Chronologically, with Qualitative Strands in Grey

Time	Data source: Class / cohort	Type of data
Mar.-Apr. '19	Many classes (pilot)	QUAN: survey ( <i>know</i> , interest)
May.-Jun. '19	Several psych. classes (pilot)	QUAN: survey (interest)
Aug.-Sep. '19	Undergrad. psych. / '19	QUAN: survey ( <i>know</i> , interest)
Aug.-Sep. '19	Other undergrads / '19	QUAN: survey ( <i>know</i> , interest)
Fall '19	Undergrad. psych. / '19	QUAN: source-evaluation, -use tasks ( <i>do</i> )
Nov. '19	Undergrad. psych. / '19	QUAN: survey ( <i>know</i> , interest)
Nov. '19	Other undergrads / '19	QUAN: survey ( <i>know</i> , interest)
Spring '20	Graduate students	QUAN: survey ( <i>know</i> )
Mar '21	Undergrad. psych. / '18	QUAL: focus-group interviews ( <i>feel</i> )
Apr. '21	Undergrad. psych. / '19	QUAN: survey ( <i>know</i> , interest)
Mar.-May '22	Undergrad. psych. / '19	QUAL: individual interviews ( <i>feel</i> )
Mar. '22	Undergrad. psych. / '19	QUAN: source evaluation task ( <i>do</i> )
Mar. '22	Undergrad. psych. / '19	QUAN: survey ( <i>know</i> , interest)
Apr. '22	Undergrad. psych. / '17, '18	QUAN: source use task ( <i>do</i> )
Jun. '22	Undergrad. psych. / '19	QUAN: source use task ( <i>do</i> )

Tashakkori et al. (2020, p. 128) have devised a typology of mixed-methods designs that provides more nuanced descriptions of complex mixed-methods studies than previous models. Their typology describes four main families of mixed-methods research designs which differ from one another as to when the strands, and their mixing, occur:

- *Parallel designs*: two or more strands (QUAL, QUAN) occur in parallel, either simultaneously or with a short time interval
- *Sequential designs*: two or more strands (QUAL, QUAN) occur chronologically, with one following the other, based on data, results, and conclusions in the previous strand
- *Conversion designs*: mixing occurs when some of the data is transformed or analyzed thematically or statistically
- *Hybrid designs*: when features of two or more of the above designs are combined, and integration of three or more strands occurs at multiple stages

This IL project contains elements of both *sequential* and *conversion* designs and therefore fits the description of a *hybrid mixed-methods design*. The *sequential* nature of the project is evident in Table 1, where QUAN and QUAL data collection was performed on multiple occasions over three years. Although not explicitly illustrated in the table, mixing occurred when each strand's data, results, and inferences informed the development of the next strand. In order to answer the research questions comprehensively, new questions were created, and new participants were recruited when necessary. The *conversion* of data occurred in the two assignment-based tasks, where qualitative, written assignments were scored quantitatively, as described briefly in the Methods section and in detail in Paper 1.

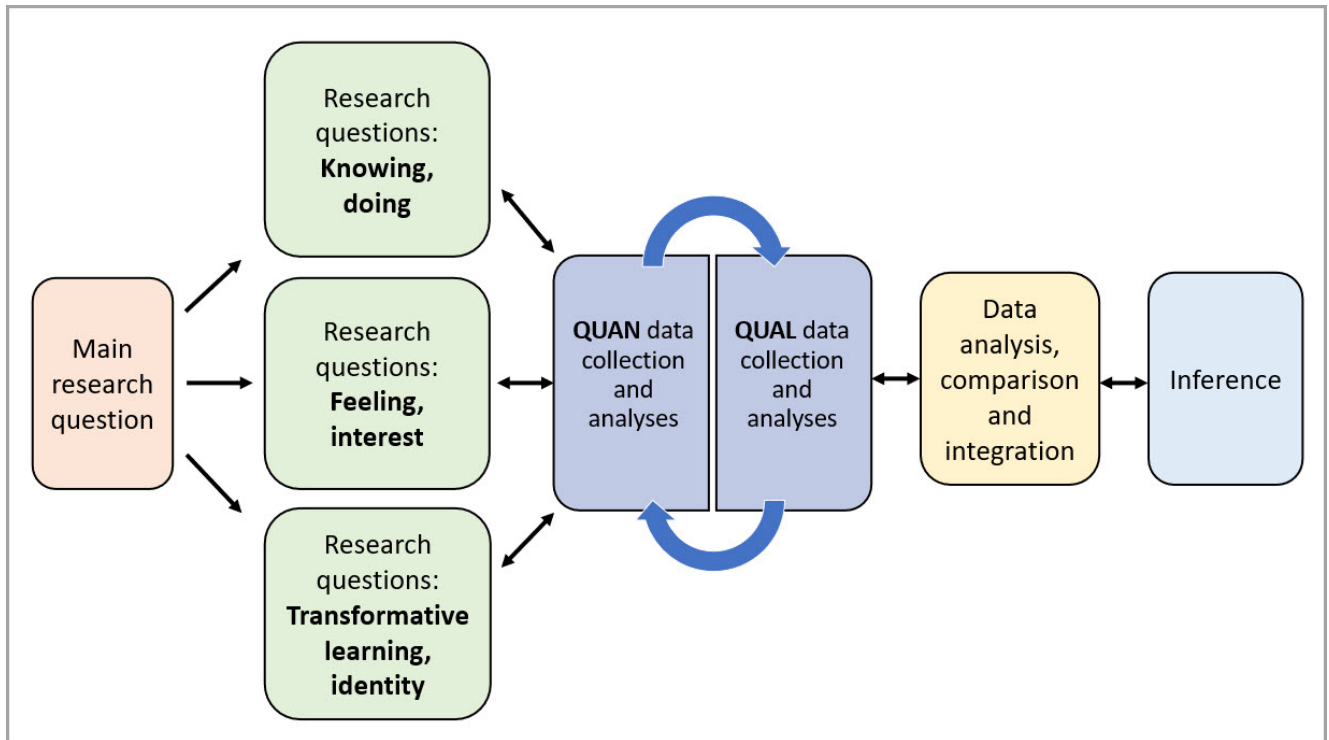
In addition to differentiating between families of mixed-methods designs in their model, Tashakkori et al. (2020, p. 127) also distinguish between three types of data sources (participant samples):

- *Multisample*: data collected from two or more groups
- *Same-/subsample*: data collected from the same sample or subsample
- *Multilevel*: QUAN and QUAL data collected at different hierarchical levels

Sampling in this project has elements of both *same-sample* and *multisample*. The psychology students that were followed over time represent a *same-sample* data source, while other participants – for example graduate students and undergraduates who were not part of this group – represent *multisample* data sources (see Table 1). The project as a whole is thereby considered *multisample* and categorized as a *multisample hybrid mixed-methods design*.

Figure 2 illustrates the pervasive mixing of QUAN and QUAL strands, the mutual influence of strands on each other, and iterative components of this research project. QUAN data collection and analyses are integrated with and influence QUAL data collection and analyses, and vice versa, at multiple stages throughout the research. Mixing occurs iteratively between strands, either in parallel or chronologically, as also seen in Table 1.

Figure 2: Conceptualization of the Mixing of Methods and their Iterative Nature



### 3 Methods for individual studies

Information about the participants, materials, and procedures for the individual studies in the PhD project is provided in Table 2. The table is divided into sections for pilot studies, studies with students that were followed over time, and other related studies. Methods, analyses, and results are described in more detail below, as well as in the individual papers.

Table 2: Materials, Measures, Participants, Data Collection

Study/Paper	Materials	Measure	Participants/Cohort	Data collection
<b>Pilots</b>	Pilot IL test	Know	Undergraduates from 2 universities	Apr. '19
	Expert evaluations	Know	4 IL experts	May '19
	Pilot interest questionnaire	Feel	Psychology undergraduates at different levels	May '19
	Think-aloud protocol	Know	3 first-year bachelor students 2 final-year upper secondary school students	June '19
<b>3-year study</b>	Survey: TILT <sup>a</sup> and TRIQ <sup>b</sup>	Know, feel	1 <sup>st</sup> , 4 <sup>th</sup> , and 6 <sup>th</sup> semester psychology students Cohort '19 (IL instruction fall '19)	Sep. '19 (sem. 1 start) Nov. '19 (sem. 1 end) Apr. '21 (sem. 4 end) Apr. '22 (sem. 6 end)
	Tasks: Source evaluation, source use	Do	1 <sup>st</sup> and 6 <sup>th</sup> semester psychology students Cohort '19	Fall '19 Spring '22
	Task: Source use	Do	6 <sup>th</sup> semester psychology students (bachelor) Cohort '17 Cohort '18	Spring '20 Spring '21
	Focus-group interviews	Feel	6 <sup>th</sup> semester psychology students (bachelor) Cohort '18	Mar. '21
	Individual interviews, Interest-o-meter	Feel	6 <sup>th</sup> semester psychology students Cohort '19	Mar.–May '22
<b>Participants who were not involved in the pilot or 3-year studies</b>				
<b>Paper 1,2</b>	Survey: TILT and TRIQ	Know, feel	First-year undergraduates from several universities	Sep. '19 (sem. 1 start) Nov. '19 (sem. 1 end)
<b>Paper 1,3</b>	TILT	Know	Master and PhD students from several universities	Spring '20
<b>Paper 3</b>	TILT + 3 questions about interest/need/effort	Know, feel	Undergraduate and PhD students from several universities	Fall '20

<sup>a</sup> TILT = Tromsø Information Literacy Test. <sup>b</sup> TRIQ = Tromsø Interest Questionnaire.

Although the individual studies in this project had different participants, one sample was followed over time. This was a group of psychology students at UiT The Arctic University of Norway (UiT) that were followed over the first three years of their undergraduate education, from fall 2019 to spring 2022. This group is referred to as the *2019-cohort* or *cohort '19*. Students came and went, possibly due to the pandemic to some extent, so few individuals in this group could be followed longitudinally over three years. This meant that the sample size for repeated measures, i.e., following the same students over time, was smaller than anticipated at the start of the doctoral project.

At UiT, some psychology students enroll in a bachelor's program, which in Norway takes three years, and others enroll in a six-year professional-studies program. These programs have the same IL instruction and IL-related learning outcomes, so data collected from students in both programs regarding their IL are comparable. Students in all psychology courses are required to use sources correctly in their writing assignments and this requirement is included in each course's assessment criteria. In their first semester, all psychology students involved in this project had an obligatory course called *Thinking, learning, and writing in higher education*, a flipped-classroom academic-writing course with embedded IL. In their third year, all students had library instruction in database searching.

Data from cohort '19 students that were followed longitudinally – the *longitudinal group* – were analyzed as paired samples. Cross-sectional data, collected from students in the 2019-cohort who were *not* possible to follow longitudinally, as well as data collected from students in cohorts that began in 2017 and 2018, were analyzed as independent samples. Students in these cross-sectional samples had the same IL instruction and assessment as students in the longitudinal group, so their data are comparable.

Four tools, described in the Methods sections of Papers 1 and 2, were used to measure various aspects of IL in this doctoral project. All four tools belong to the *Tromsø Information Literacy Suite* (TROILS), and are freely available for others to use on the TROILS website (<https://site.uit.no/troils>). The development, testing, and preliminary results of three of these measures – the IL-knowledge test, the source-evaluation measure, and the source-use measure – are described in Paper 1. The development, testing, and preliminary results of the fourth measure – the interest questionnaire – is described in Paper 2. Data from these measures were collected from students in the 2019-cohort at four points in time: twice during their first semester – before IL instruction and at the end of the semester – and toward the end of their

fourth and sixth semesters. Since there were very few respondents in the fourth semester, these data are neither included in the individual studies nor in the four papers.

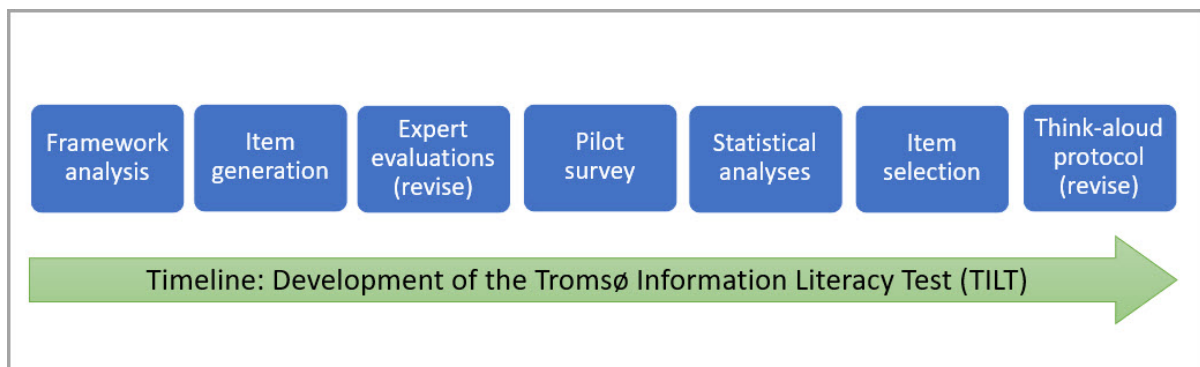
### 3.1 Methods: Paper 1

Paper 1 describes the development and use of three tools for measuring IL *knowing* and *doing*. This paper also addresses the dimensionality of the IL construct.

#### 3.1.1 Information literacy knowledge test: *Knowing*

To measure students' knowledge of key aspects of IL, we developed a 21-item test, later dubbed the *Tromsø Information Literacy Test (TILT)*. The pilot (in Norwegian) and final versions (in English and Norwegian) of TILT are found in Appendix 6. Figure 3 illustrates a chronology of test development. Methods involved are described briefly below and in detail in Paper 1.

Figure 3: Timeline: Development of the Tromsø Information Literacy Test (TILT)



TILT's psychometric properties were documented in multiple ways during item generation and item selection to assure its validity<sup>6</sup> and reliability in assessing IL knowledge. A framework analysis (see Appendix 1) contributed to the validity of the IL test. Nine IL frameworks were compared with the aim of finding the one that best fulfilled our criteria,

---

<sup>6</sup> In my descriptions of the measures' psychometric properties, I frequently use the term *validity*, instead of subcategories such as content, construct, or discriminant validity. This choice is based on Streiner and Kottner's (2014, p. 1976) argument that validity is a unitary concept, namely the extent to which an instrument actually measures the construct it is intended to measure. They write that "since the 1970s, all validity testing is seen as a form of construct validation. We no longer speak of different types of validity, such as criterion validity, discriminant validity, convergent validity, or the myriad other 'types' that have been introduced over the years."

namely that it should (1) encompass central facets of IL, thereby assuring the validity of test items in representing the IL construct; (2) specify learning outcomes, to make item generation more straightforward; and (3) apply to most disciplines in HE. The chosen framework was the *Australian and New Zealand Information Literacy Framework* (see Paper 1).

Item generation for the IL-knowledge test was informed by the framework's learning outcomes. After generating a pool of 50 multiple-choice items, four IL experts evaluated the items for validity (clarity, content accuracy, and objectivity and bias), and items were revised based on the experts' feedback. A pilot test was then conducted with undergraduates from several Norwegian universities. Results were psychometrically analyzed (see e.g., Irwing et al., 2018) to select those items that had a suitable range of difficulty (item difficulty) and that correlated positively to total test scores (item discrimination). The items remaining after this selection process all related to finding, evaluating, and using sources – three key, source-related facets of IL. For these and other reasons (see Paper 1), I have chosen to focus on assessing these three source-related aspects of IL in my research. By limiting the scope of my research in this manner, it is not my intention to imply that IL consists *solely* of these specific facets, only that they are central to the IL construct. In order to have an equal number of items in these three facets, four additional items were then added, although these had neither been tested in the pilot nor rated by experts – a limitation in this study.

Exploratory factor analyses (see Paper 1, Appendix A) were performed to explore whether extracted factors aligned with the three facets of IL in the test. If this had been the case, unrelated items could be eliminated. However, no interpretable, meaningful factors were found so there was no useful basis for further item selection. Results of the factor analyses were also used to explore the dimensionality of the IL construct, i.e., whether it is a unidimensional or a multifaceted construct.

After item selection was completed, think-aloud protocols were performed with five students in the target group to test items for readability and comprehension.

The final 21-item IL-knowledge test TILT is found in Appendix 6 (in English and Norwegian), together with the pilot version (in Norwegian). TILT was employed in the rest of the project as part of a survey that also included (a) an interest questionnaire (see Paper 2), (b) demographics, such as gender, age, and level of HE, and (c) a consent form (see Appendix 4). This survey was distributed on the online survey software Qualtrics to undergraduates at

several Norwegian HE institutions, including psychology students at UiT, at the beginning and end of their first semester. In addition, TILT's English translation was distributed to graduate students at several international universities to compare with results from undergraduates and Norwegian samples. Results are included in Paper 1.

Evidence of TILT's validity was supported by analyses calculating the degree to which the test distinguished between students at different education levels (bachelor's, master's, PhD), languages (Norwegian or English), and between test scores before and after IL instruction. TILT's reliability was checked with a test-retest in undergraduates (temporal consistency), and was found sufficiently reliable (see Paper 1, pp. 88-89).

### **3.1.2 Source-evaluation and source-use measures: *Doing***

Curricula for psychology courses at UiT contain learning outcomes relevant to IL, including the ability to critically evaluate and correctly cite information sources in their written work. I chose to use mandatory authentic assignments from first and sixth semester courses for measuring IL skills in this project, after modifying them slightly for assessment purposes in cooperation with faculty teaching these courses. After receiving approval from NSD, and with the consent of the participants, I collected and analyzed data from students' assignments and linked them to their survey data for *know* and *feel* measures. Each student's data from the separate measures was anonymized before assessing assignments.

IL *doing* was assessed in students' term papers in the first semester (Paper 1) and bachelor's theses in the sixth semester (Paper 4). Two measures were employed. (1) In the source-evaluation measure, students earn points for each source-evaluation criterion – such as authority or accuracy – they list in an annotated bibliography of sources they have chosen to use in their papers/theses. (2) In the source-use measure, their papers/theses are graded on a rubric for assessing reference technique and use of sources. These assignment-based *do* measures are described in detail in Appendix 7. Both measures involve assessing textual data from actual writing assignments. The transformation of this qualitative, narrative data to quantitative, numerical data (scores) represents a mixing of methods, what Tashakkori et al. (2020, p. 129) refer to as a conversion mixed-methods design.

Both measures arguably have inherent validity since they were designed by teachers, or by librarians in cooperation with teachers, partially to assess students' IL skills. When there were



several raters, interrater reliability was calculated and found satisfactory (see Paper 1, p. 88, and Paper 4, p. 7).

The relationship between *knowing* and *doing* was explored with bivariate correlation analyses to determine whether correlations between TILT scores and *do*-measure scores were significant. In addition, multiple regression was used to help determine whether TILT scores and amount of HE (predictor variables) could predict source-evaluation and source-use scores (outcome variables). See Paper 1, p. 90, for details of these analyses.

### **3.2 Methods: Paper 2**

While Paper 1 describes the development, testing, and use of the IL *know* measure, Paper 2 describes the same for the IL *feel* measure. This is a tool for measuring interest, informed by Hidi and Renninger's (2006) four-phase model of interest development. No one had previously created a self-report measure that could identify distinctions between the phases in their model. We therefore designed a measure to do this, and tested it with a specific object of interest, namely IL. In order to operationalize Hidi and Renninger's (2006) model, we designed a theoretically grounded, domain-tailorable self-report measure dubbed the *Tromsø Interest Questionnaire* (TRIQ; see Appendix 8). Before this study, no existing self-report measures of interest could identify distinctions between the model's four phases.

Hidi and Renninger's model describes variation between interest phases using four key variables – situation dependence, positive affect, competence, and meaningfulness. Each phase has a unique psychological architecture based on differing levels of these variables. With TRIQ, we divided the competence construct into two (competence level and competence aspiration) and added a measure for self-regulation and for general interest. The final subscales, each containing 2-6 items, are: General interest, Situation dependence, Positive affect, Competence level, Competence aspirations, Meaningfulness, and Self-regulation.<sup>7</sup>

In Paper 2, we report how the quality of interest – as measured by subscales – varies by interest phase. Two analyses that demonstrate this phase distinction are (1) multivariate analyses of variance (MANOVAs), that show differences between each subscale score by

---

<sup>7</sup> Since Self-regulation extends beyond Hidi and Renninger's theoretical grounding, we did not focus on this variable in the doctoral project.

phase, and (2) discriminant function analyses, which do the opposite, namely identify distinctions between the phases, both conceptually and through peoples' reported interest experiences, based on subscale scores. In other words, they demonstrated how a person's interest is concretely linked to their subscale scores.

In Paper 2, these analyses were employed in two studies: Study 1, with *self-chosen* objects of interest, and Study 2, with a *specified* object of interest, namely students' interest in being or becoming information literate, which in this doctoral project denotes IL *feeling*. Analyses with descriptive statistics comparing results from Study 1 and Study 2 provide further evidence of TRIQ's validity and robustness. The object of interest being measured did not introduce noteworthy variance in the findings, demonstrating that TRIQ can be used with many different objects of interest.

Evidence of TRIQ's reliability was provided by testing for (1) the *internal consistency* of its subscales, measured with Cronbach's alpha, and (2) its *temporal consistency*, measured with paired-sample test-retest t-tests and ICC by phase (Paper 2, p. 4).

### **3.3 Methods: Paper 3**

Paper 3 is a cross-sectional study. The aim of this study is to determine how self-aware students are of their IL knowledge levels, i.e., their metacognition. How accurate are they at estimating what they know? In the first two papers, students' IL levels were assessed using objective *know*, *do*, and *feel* measures, while Paper 3 also explores students' *perceptions* of their IL levels. Participants were 760 students from several disciplines at HE institutions in nine countries, at three HE levels: undergraduate (first-semester), master's, and PhD.

Participants were recruited via e-mail and social media. Their survey consisted of demographics, a definition of IL, and the IL-knowledge test TILT (in Norwegian and English). Before taking the test, they were asked to estimate – given the IL definition provided – how many of the 21 questions they thought they would answer correctly. After taking the test, and before receiving their score, they were asked how many questions they thought they had answered correctly. Some of the undergraduate and PhD students were also asked to rate, on a Likert scale of 1-6, (a) their interest in being or becoming an information literate person (*interest*), (b) their perceived need to learn more IL skills (*need*), and (c) how much effort they would make, knowing themselves, to develop stronger IL skills (*effort*).

Depending on their test scores, participants at each HE level (undergraduate, master's, PhD) were divided into groups of low and high performers, using a median split, to examine how these 2 groups differed in their score predictions. Two-way ANOVAs, with performance level and HE level as independent variables, were performed to determine whether these two variables, or their interaction, had a main effect on the accuracy of students' before- or after-test accuracy scores. Additional two-way ANOVAs, with HE level and gender as independent variables, were performed to determine whether these two variables, or their interaction, had a main effect on either IL test scores or the accuracy of students' before- or after-test accuracy scores. Correlations, with Pearson's  $r$ , were calculated for test questions about *interest*, *need*, and *effort*.

Not all results from analyses are reported in the published version of Paper 3 because the journal has word limits and does not publish supplementary materials. However, two appendices after Paper 3 in this document provide additional statistics. The first appendix includes descriptive statistics and statistical analyses regarding students' estimated scores on TILT, and the second contains descriptive statistics for questions about interest, need, and effort.

### **3.4 Methods: Paper 4**

Paper 4 is the culmination of this doctoral project, with the goal of understanding the development of IL in students over the first three years of an undergraduate education. The project focuses on three major, source-related facets of IL – finding, evaluating, and using information – which are represented in most IL definitions and frameworks and are generally recognized as key aspects of IL by its researchers and practitioners.

Paper 4's study uses mixed methods to explore the development of IL *knowing*, *doing*, and *feeling*, and the associations among them, over time. It includes results from the longitudinal and cross-sectional samples. All participants had comparable IL instruction and assessment. Although data was collected at four times, only data from the beginning and end of the three-year period are analyzed and presented in Paper 4. The amount and types of analyses performed were limited due to small sample sizes and article-length issues. (It can be a challenge to comply to word limits in mixed-method articles, as all methods must be described in detail.)

Quantitatively, *knowing* was measured with the IL-knowledge test TILT, and *doing* was measured with authentic assignments that involved evaluating and using sources. *Feeling* was quantified with the interest questionnaire TRIQ, as *feeling* is conceptualized mainly by interest (in being or becoming information literate) in this project. In addition to the QUAN measures, QUAL measures – namely focus-group interviews and in-depth, individual interviews – were conducted in order to provide students with more of an opportunity to reflect on their development as information-literate people than was possible in the survey. By mixing QUAN and QUAL methods, we could better understand the totality of students' IL development and more deeply explore (1) any changes in identity they may have experienced as a result of becoming more information literate, and (2) how their interest in being or becoming information literate changed over time and how this related to *knowing* and *doing*.

### **3.4.1 Quantitative methods**

Sixth-semester UiT psychology students, in bachelor's and professional-studies programs, completed the survey that includes the *know* and *feel* measures TILT and TRIQ. Samples of sixth-semester students in the *do* measures varied, as those in the bachelor's program had assignments appropriate for both measures – an annotated bibliography (source evaluation) and their bachelor's theses (source use) – whereas those in the professional-studies program had neither. To increase sample size, sixth-semester students from previous comparable cohorts – those beginning in 2017 and 2018 – were included in cross-sectional samples, and their bachelor's theses were assessed for source use. There were no equivalent source-evaluation assignments from previous cohorts, however, so the sample size for this *do* measure was low. Although the source-use rubric is not used by teachers in the assessment of bachelor's theses, I employed it in this study to make results comparable to assessments from their first year. Source-use assessments were based on the first six pages of bachelor's theses, approximately the same number of pages as in first-year student papers.

In terms of the reliability of *do*-measure assessments in the third year, I performed the assessments for the source-evaluation measure alone, and since I was a qualified rater in the first year, these assessments can be considered reliable. For the source-use measure, another rater assisted in the assessments, and interrater reliability, as measured with weighted Cohen's kappa, was sufficient (0.77, see Paper 4).

### **3.4.2 Qualitative methods**

After collecting and analyzing QUAN data from the four measures described above, our curiosity took us beyond what we were able to answer with the QUAN methods alone. The multiple-choice answer alternatives in TILT and Likert scales in TRIQ were closed-ended and did not offer respondents an opportunity to elaborate on their IL development and their identities as information-literate people. We chose therefore to complement the QUAN methods with QUAL methods, namely interviews, that enabled us to increase our understanding of these issues and gain insight into students' perception of their growth as information-literate individuals. Students in their sixth semester were most suitable for interviews for several reasons: (a) they were mature students and could reflect on their IL development over the past three years; (b) they were least represented in the QUAN data and there was therefore a need and desire to supplement these data; and (c) those in the bachelor's program were currently working on their bachelor's theses, where IL skills were relevant and essential.

I chose to first conduct focus-group interviews (FGI's) because of their advantageous group effects. Participants can mutually stimulate each other in focus groups, leading to increased interaction and engagement in the topic and providing the sense of security needed for participants to share their thoughts (Gawlik, 2018). FGI's are appropriate for exploratory studies, as varied – and often conflicting – viewpoints are expressed and discussed (Brinkmann & Kvale, 2015, p. 175).

As the project progressed and the focus on transformative learning and identity increased, I decided to conduct a second round of interviews with questions informed by these theoretical lenses. These were individual interviews, allowing each student more opportunity than in focus groups to elaborate on their personal IL development in the three years since the start of their undergraduate education.

All interviews were recorded with audio rather than video since facial expression and body language were deemed less significant in this context. All interviews were semistructured, designed to allow participants the opportunity to introduce new topics to the conversation. This could be advantageous if informants had perspectives that were not specifically covered in the interview guide or anecdotes they wished to share.

At the start of the interviews, after presenting myself and my research, I asked informants to say a few words about themselves to break the ice. I then provided instructions about the procedure, information about their privacy, and the definition of IL used in this project. Each informant answered every question. Follow-up questions – some pre-composed and others spontaneous – were posed when deemed necessary or fruitful. At the end of the interview, I asked if informants had anything else to bring up or any experiences to share, providing them with the opportunity to introduce topics relevant to IL, in an open-ended manner, that had not previously been addressed.

In individual interviews, several questions in the interview guide related to the student's perceived identity as an information literate person. However, after reading through transcripts of these interviews, my co-supervisor Mariann Solberg and I recognized that the data regarding identity were not saturated. We therefore conducted follow-up interviews with those informants from the individual interviews who we agreed had shown evidence of a change in identity. These follow-up interviews were conducted by Mariann, hoping that the variation provided by our two styles of questioning would contribute to discovering more about the students' identity changes, transformative learning, IL development, and interest in being or becoming information literate.

In this second round of individual interviews, a new, mixed-methods technique was implemented. Students were presented with an empty graph, where the x-axis represented time, in semesters, from the start of their undergraduate education until the present, and the y-axis depicted their level of interest in being or becoming information literate, from very low to very high. In this so-called *interest-o-meter* – based on the *feelometer* instrument originally used by Hetland and Vittersø (2012) – participants were asked to draw a line/curve depicting changes in this interest level over time, and to provide verbal descriptions simultaneously. Their graphs are provided in Appendix 9.

Shortly after conducting interviews, recordings were transcribed. Concepts, themes, and ideas were deemed more central in this study than details of language, so transcriptions were simple, capturing words and emphasis, rather than detailing linguistic annotations, dialects, and emotional expressions, like sighing and laughter, which are more typical of conversation and discourse analyses (Kvale, 2007). Note was made during interviews when pronounced feelings were displayed. Participants, procedures, and materials in the focus-group and individual interviews are described below.

### **3.4.3 Focus-group interviews**

Participants for the FGI's were recruited from third-year students in psychology who had begun their bachelor's study in 2018 and were currently writing a bachelor's thesis. Although the 2018-cohort was not included in the longitudinal study, they had the same core curriculum and IL instruction as the 2019 longitudinal cohort, so their IL learning experience was analogous. Students were recruited via the class's learning management system and awarded a gift card for participation. Nine psychology students (7 women, 2 men) volunteered and were divided into three groups of three, with the men in different groups. Because of the pandemic, the FGI's were conducted digitally.

The interview guide (see Appendix 10) includes main questions and follow-up questions. These focus primarily on students' self-perceptions of their IL development, including their knowledge, skills, and attitudes.

### **3.4.4 Individual interviews**

Participants for individual interviews were recruited from the 2019-cohort of psychology students, who were now in their sixth semester. Students in the bachelor's program were in their final semester and writing their bachelor's theses, while those in the professional-studies program were halfway through their six-year program. Four students volunteered for individual interviews, one woman and one man from the bachelor's program, and one woman and one man from the professional-studies program.

Individual interviews were conducted physically in spring 2022 – as pandemic restrictions had by then been eased – and recorded with audio. The three follow-up interviews were performed by my co-supervisor, as previously mentioned.

The interview guides for both rounds of individual interviews are found in Appendix 10. The first round contained some of the same questions as the FGI's, plus new questions focusing on identity and TL. The second round contained more in-depth questions on these themes, as well as questions related to the interest subscales from TRIQ (see Section 2.5.4). Second-round informants also used the *interest-o-meter* to illustrate changes in their interest in being or becoming information literate over the past three years, describing the highs, lows, and overall evolution of their interest (see Appendix 9). The goal of the individual interviews was to understand the totality of students' IL development more thoroughly, including the degree

to which being information literate is part of their identity and whether transformative learning with regard to IL has occurred.

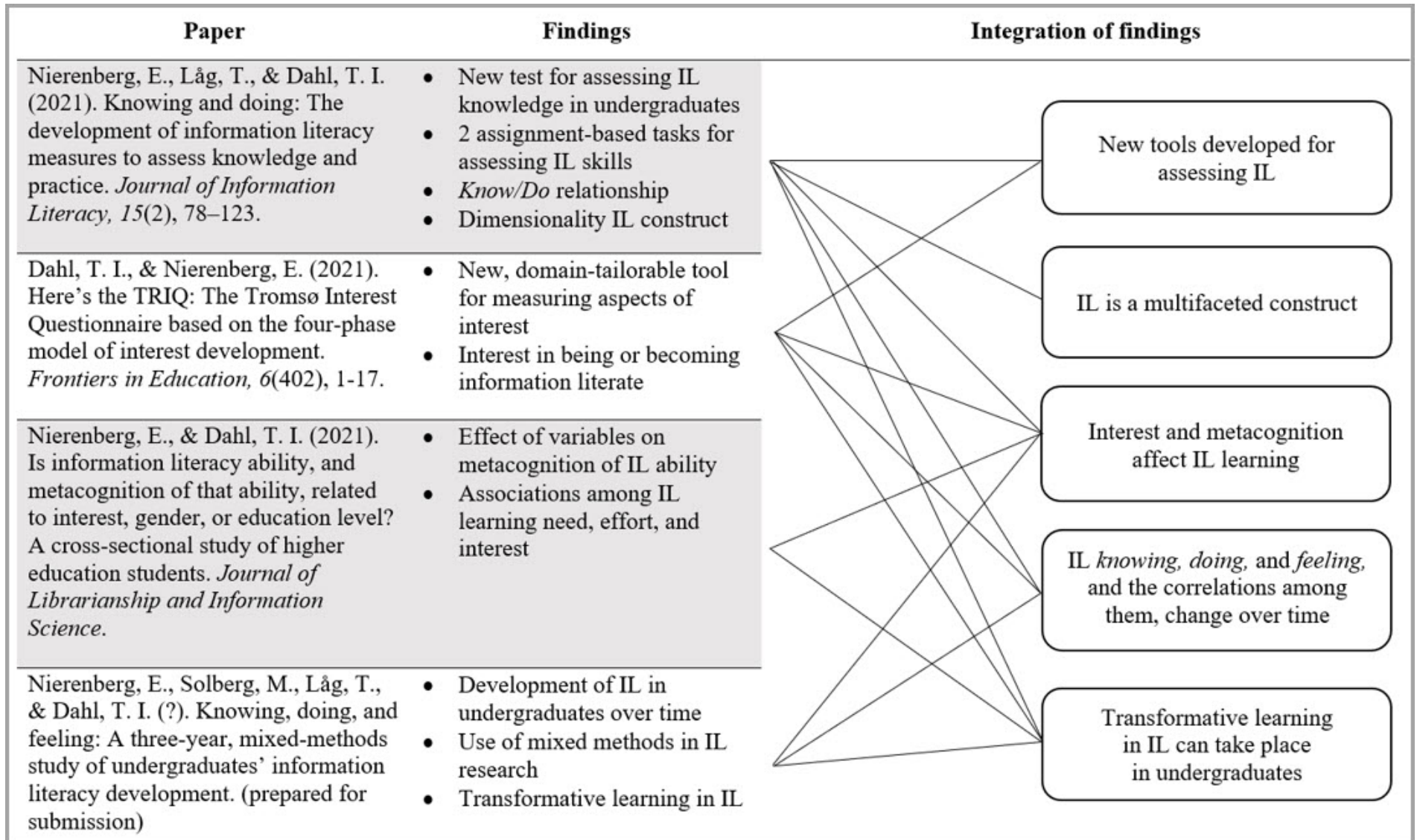
In order to identify and describe central concepts, ideas, and themes from the interviews, my co-supervisor and I conducted a thematic analysis of the transcriptions. We did this together since each of us had conducted a round of individual interviews, so there were two judges involved in analyzing the interview data. The QUAL data analysis software NVivo was used for coding. Braun and Clarke's (2006, 2012) six-phase method of thematic analysis was employed to analyze the data. First, we familiarized ourselves with the data by reading through the transcriptions. Second, we searched for initial codes in the transcriptions. The creation of codes was based both on describing (a) previously identified topics and concepts from theory, a deductive, top-down approach, and (b) the content of the data themselves, an inductive, bottom-up approach (Braun & Clarke, 2012). Coding was an iterative process; as new codes were created and refined, several previously assigned codes were changed and recoded. Third, potential themes in the data were constructed by combining codes that had similarities, and relevant data for each theme were gathered. The themes were relevant to the main research questions and represented a consistent and meaningful pattern in the dataset. Some themes were related to each other, while others stood alone. While the creation of codes was a descriptive process, the creation of themes was more analytic. Fourth, themes were reviewed to check if they corresponded to the data and were supported by a sufficient amount of data, and that the data were saturated. Fifth, themes were defined and named, specifying what is unique about each. Sixth, a scholarly report of the analysis was written up (see Paper 4), including examples of quotes from interviews to illustrate themes.

### **3.5 Coherence**

In this section, I explain how the papers in the dissertation fit together as whole. I demonstrate the significance of each of the papers and how the findings cohesively address the main research question. Figure 4 maps the relationship between the four papers, and how they contribute to answering research questions. Lines connect key findings from each paper to the main findings of the overall dissertation.



Figure 4: Integration of Findings Across Articles



In Papers 1 and 2, instruments were designed, tested, and employed to measure IL knowledge (*knowing*), skills (*doing*), and attitudes (*feeling*). These tools were used in the remainder of the project – as illustrated in Figure 4 – to measure students’ metacognitive awareness of their IL abilities (Paper 3) and the development over time of IL *knowing*, *doing*, and *feeling* in psychology undergraduates (Paper 4). Papers 1 and 2 thereby provide the tools necessary to answer the overarching research question of how students develop as information-literate individuals in the first three years of an undergraduate education.

Paper 1 addresses (1) the assessment of IL knowledge and skills, and also explores (2) the relationship between what students know and what they do in practice, which is also related to the overarching research question, and (3) whether the IL construct is unidimensional or multifaceted. This third point is controversial in the IL community (see Paper 1, Section 4.3). It surfaces in many contexts, for example when developing IL tests, and it is therefore relevant to the project as a whole.

Development of the interest questionnaire TRIQ is described in Paper 2. TRIQ was created in two versions in order to control for reliability and validity, and similar patterns were found in both versions. The first version measures interest in a self-chosen object of interest, and the second measures interest in a specified object of interest, namely students’ interest in being or becoming information-literate people. Quantitatively, in other words, *feel* is conceptualized mainly by interest in this project. In the remainder of the project, this second version of TRIQ was employed to measure phases and aspects of this specified interest in students to examine how it changes over time. So, while Paper 1 documents the development of tools for measuring IL *knowing* and *doing*, Paper 2 documents the development of a tool for measuring IL *feeling*.

In Paper 3, students’ self-awareness of their IL abilities (metacognition) was measured using the IL-knowledge test developed in Paper 1. Becoming metacognitively aware, i.e., acquiring the ability to accurately assess what we know, may lead to an increase in learning or task performance (Kruger & Dunning, 1999). Metacognitive ability thereby represents an important part of learning, also IL learning. It was therefore important in this project to determine which factors were related to students’ metacognition of their performance on the test, including whether the act of taking the test itself was related to metacognitive accuracy. Associations among interest, need, and effort (see Section 3.3) help us explore the role of interest in motivating learning, and are therefore relevant to students’ IL development. If our

goal as IL teachers is for students to exert effort in pursuing IL learning, it is important to know how and when interest and need play a role in motivating learning, alone and together. Students in the longitudinal and cross-sectional samples did not answer questions about interest/need/effort in their first semester, so it was not possible to study their growth over time. Therefore, results over time are neither reported in Paper 4 nor in this extended summary.

Paper 4 employs mixed methods to follow the undergraduates' *knowing, doing, and feeling* – and the relationship between them – over three years. This paper thereby directly addresses the overarching research question in this project, and employs tools developed in the first study (Papers 1 and 2). In addition to these QUAN measures, focus-group and individual interviews were conducted to obtain a more nuanced description of the students' development, as well as their identities as information-literate people and the degree to which transformative learning takes place.

## 4 Results: A short summary

### 4.1 Quantitative findings

In this section I present QUAN findings, and connections between them, that were not explicitly reported in the papers, although some of the results may have appeared in tables.

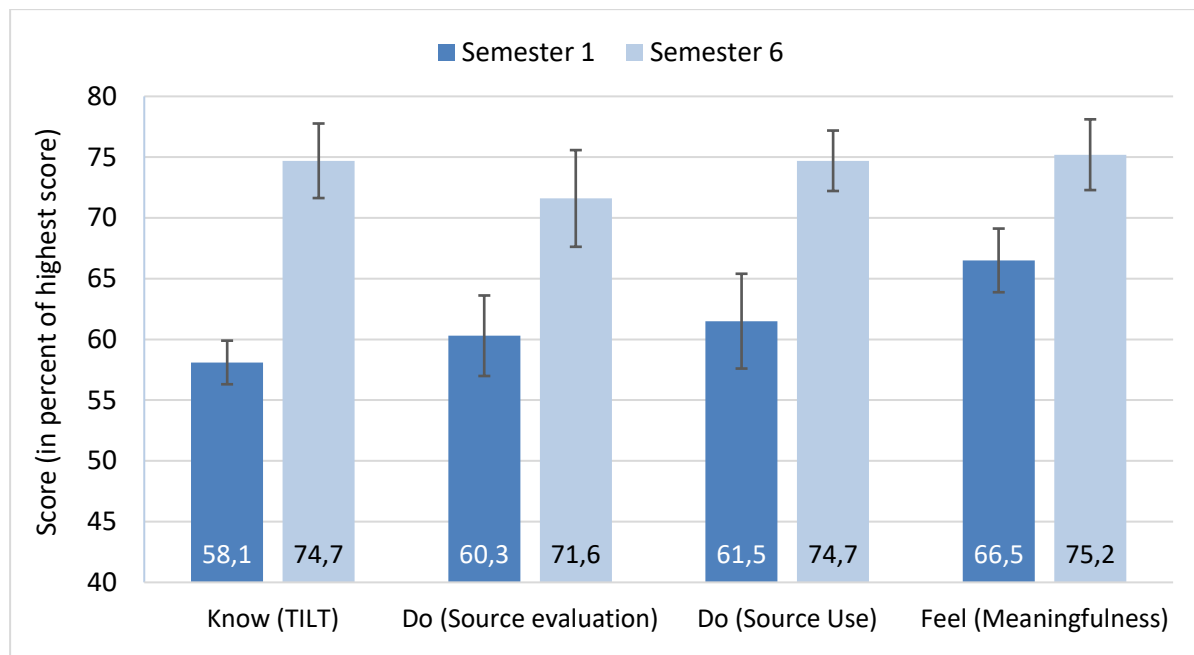
Paired and cross-sectional samples. The four papers in this dissertation report QUAN results from several different studies, each with its own sample(s) of participants. All papers include data from undergraduate psychology students at UiT that began in fall 2019, some of whom were possible to follow over three years. Thirty-three of these students filled out the survey (with TILT and TRIQ) at the beginning of their first semester and the end of their sixth semester. Data from these students comprised the *paired sample*. Data that were only available from one of those semesters were included in *cross-sectional samples*. Also included in these cross-sectional samples are sixth-semester data from the paired sample. This is justifiable because no data from the same student are duplicated in cross-sectional samples, as their first-semester data are not included. To best generalize QUAN findings from the entire project, *knowing*, *doing*, and *feeling* findings presented in this section are taken from these cross-sectional samples.

Know, do, and feel over time. Figure 5 shows first- and sixth-semester results, in percent of highest possible/achieved score, for *know*, *do*, and *feel* measures. In this figure, *know* values are based on TILT scores, *do* values are based on the source-evaluation and source-use measure scores, and *feel* values are based on the TRIQ subscale score for Meaningfulness.<sup>8</sup> All four show statistically significant growth between semesters 1 and 6. These results show the same trends as the individual studies, despite differences in actual statistics resulting from the characteristics of each study's specific sample.

---

<sup>8</sup> Since each subscale measures something qualitatively different, there is no value in combining their data in an attempt to measure overall IL interest. Although no individual subscale sufficiently represents IL interest on its own, one was chosen in order to explore relationships between *knowing*, *doing*, and *feeling*. The rationales behind choosing the Meaningfulness subscale for this purpose are found in Paper 4 - Results. If the Meaningfulness subscale provides a good indication of students' motivation to learn, as we believe it may, this is useful information to IL educators.

Figure 5: Mean Scores on Know, Do, and Feel (Meaningfulness Subscale) Measures, with 95% Confidence Intervals. Cross-Sectional Sample, Semesters 1 and 6

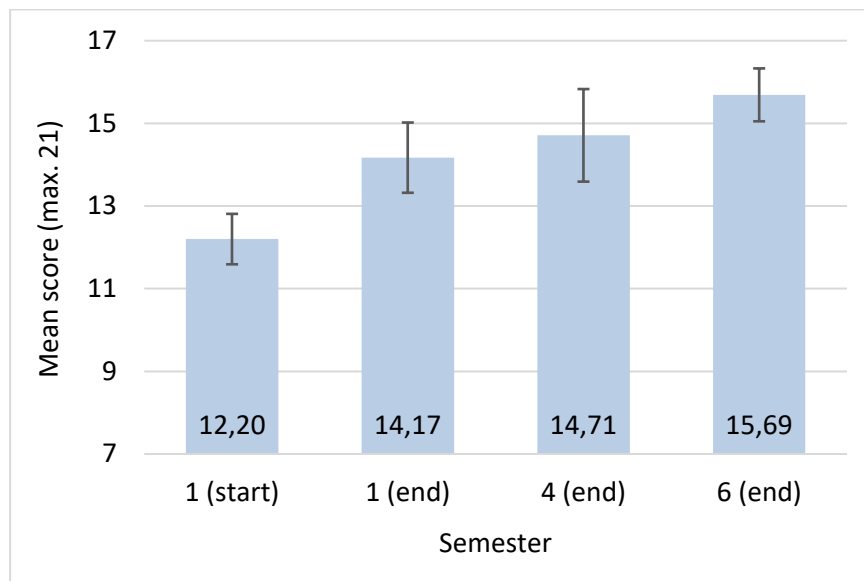


TILT scores at four points in time. Although Figure 5 shows results only from the beginning of semester 1 and the end of semester 6 for IL *knowing*, TILT was also distributed to students in the 2019-cohort at the end of their first and fourth semesters.<sup>9</sup> Figure 6 includes results from all four data collections in cross-sectional samples, to illustrate that mean TILT scores increased between each. The largest increase, nearly 2 points (of 21 possible), occurred between the beginning and the end of the first semester, before and after the academic writing course with embedded IL. The next largest increase, between the fourth and sixth semesters, was just under 1 point.

---

<sup>9</sup> Results from the beginning and end of the first semester are reported in Paper 1. Results comparing the beginning of the first semester to the end of the sixth semester are reported in Paper 4.

Figure 6: Mean TILT Scores, with 95% Confidence Intervals



TRIQ interest phases and subscale scores. In this paragraph, I explore changes in interest phases and subscale scores (see Paper 4, Table 1) in the cross-sectional data. Although individual students' self-reported interest phases changed over time, the mean for all participants remained the same; there was no change in mean interest phase between semester 1 and 6. Based on the four-phase model of interest development, this implies that no substantial changes in subscale scores should be observed. This, however, is not what we found. Subscale scores for Competency level and Meaningfulness increased significantly with time, while Competency aspirations decreased significantly. Non-significant changes include an increase in Situation-dependence, and a decrease in General interest and Positive affect.<sup>10</sup>

Metacognition. Results over time for metacognition are not presented in Paper 4 because of article-length issues, but two unexpected results – based on students' TILT-score estimates made *after* taking the test – are briefly presented here. First, for the cross-sectional sample as a whole, metacognitive accuracy *decreases* over time; students underestimated their TILT scores slightly in semester 1 and underestimated them more in semester 6. Second, in the

---

<sup>10</sup> The same trends with subscale scores were observed in the paired sample, which has approximately half the number of respondents. However, in the paired sample, mean interest phase decreased with time (see Paper 4, Table 2), which makes interpretation, well ... interesting.

sixth semester, both low and high performers underestimated their scores. So, although there was evidence of the Dunning-Kruger effect (where low-performers *overestimate* their abilities and high-performers *underestimate* them) in the first semester, this was not the case in the sixth semester. However, since confidence intervals were large and sample sizes were small, these results have low statistical power. Nevertheless, these findings – found in both the cross-sectional and paired data – are worthy of future investigation.

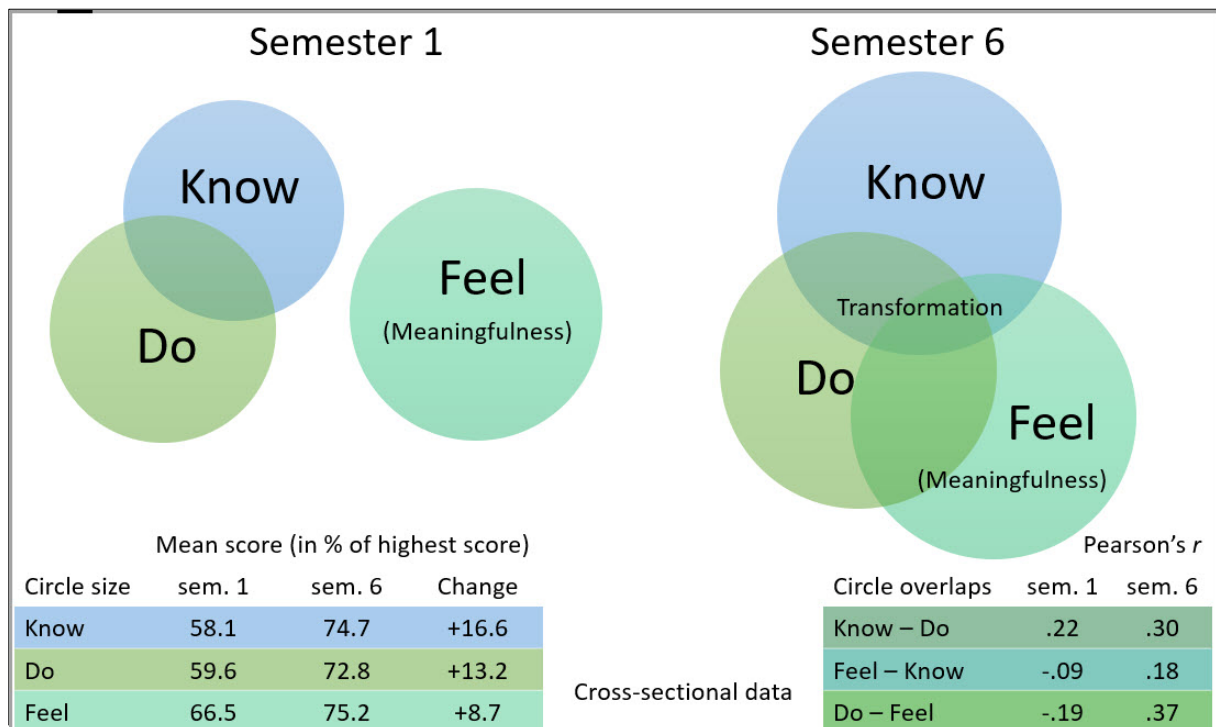
Relationships among *knowing*, *doing*, and *feeling*. Paper 1 reported results relevant to relationships between *knowing* and *doing*, showing some significant correlations between TILT scores and *do* measures in the first semester. Paper 4 reported results from the cross-sectional data, showing an increase in correlations (*know-do*, *feel-know*, *do-feel*) between the first and sixth semester. Neither the correlations themselves nor the growth in correlations over time were statistically significant. There were, however, other significant correlations found among aspects of *know*, *do*, and *feel* that were not reported in Paper 4. These are mentioned briefly below, but without details of the statistics, as new empirical data is not presented in this extended summary. There were positive, significant correlations between:

- TILT subscore for the seven source-use items and the source-use assignment score, in both semesters (*know/do*). This correlation increases from semester 1 (moderate) to semester 6 (strong), indicating integration over time of source-use knowledge and skills. This also provides evidence of the validity of the source-use items in TILT.
- TILT total scores and source-use assignment scores (*know/do*). This correlation was significant and strong in the first semester, confirming results from Paper 1, but was not significant in the sixth semester, although the correlation coefficient (*r*) was larger. This is arguably a result of the smaller sample size in the semester 6.
- source-evaluation assignment scores and TRIQ Meaningfulness subscale scores (*do/feel*). This correlation was strong in the sixth semester, but not in the first semester, showing a growing integration over time of students' source-evaluation skills and attitudes (as measured by the Meaningfulness subscale).
- average scores of *do* measures (in percent of highest score) and students' reported phase of interest in being or becoming an information literate person (1 lowest - 4 highest) in the sixth semester (*do/feel*). This correlation is negative in the first semester, but positive in the sixth semester, indicating growth over time in the association between student interest and what they do in practice with their IL.

- students' reported phase of interest in being or becoming an information literate person and the TRIQ Meaningfulness subscale score (*feel/feel*). This supports the claim that the Meaningfulness subscale is a good indicator of this interest. The correlation increases from moderate to strong from the first to the sixth semester.

Summary of quantitative findings. In addition to the significant increases in scores on *know*, *do*, and *feel* measures from semesters 1 to 6, bivariate correlations between measures also increased with time, although not significantly. Figure 7 illustrates, for the cross-sectional data from Paper 4, the growth in IL knowledge, skills, and meaningfulness over time. Actual values are shown in the tables at the bottom of the figure. Circle sizes and circle overlaps are a graphical illustration of these values, intended to portray growth qualitatively.

Figure 7: Relationships Between Know, Do, and Feel in Semesters 1 and 6: Cross-Sectional Data



Circle diameters in Figure 7 represent scores on the *know*, *do*, and *feel* measures, in percentages of the highest score. Mean scores are shown in the table on the bottom left of the figure. The *do* score is the average of the scores on the source-evaluation and source-use assignments. *Know*, *do*, and *feel* scores increase significantly from the first to the sixth semester. *Know* has the largest increase, followed by *do*.

Circle overlaps represent Pearson's *r* correlations between *know*, *do*, and *feel* scores. Correlation coefficients are shown in the table in the lower right of the figure. Negative



correlations between *feel-know* and *do-feel* in the first semester indicate that early in their studies, there is little integration between what students feel and what they know or do. In the sixth semester all correlations are positive, showing tighter integration of *knowing*, *doing*, and *feeling* with time. All correlations increased between the first and sixth semesters, with the largest increase between *do* and *feel*. However, as reported in Paper 4, the six correlations in the table are small, and none are significantly different from zero. The growth of these correlations – as illustrated by circle overlaps for the sixth semester in Figure 7 – is where I imagine transformative learning can occur (see Section 5.3).

## 4.2 Qualitative findings

### 4.2.1 Views of IL and of being information literate

In interviews, informants expressed their views of IL and of being information literate. Interestingly, some of the personal qualities that they associate with being information literate have negative connotations. Alphabetically, adjectives they used include argumentative, confident, critical, curious, hard-working, irritating, judgemental, objective, pompous, professional, reflective, skeptical, smart, superior, thorough, and trustworthy. Informants depicted being information literate as an irreversible condition, and pointed out that there are substantial risks of *not* being information literate. IL skills were described as important, instrumental, meaningful, necessary, useful, and/or valuable. Informants expressed feelings of competency, engagement, excitement, insecurity, integrity, pride, and/or satisfaction while learning IL. They felt that the process of learning IL was challenging, enlightening, overwhelming, rewarding, and/or tiresome. While learning IL, they experienced feelings ranging from ignorance to frustration to mastery, and many still felt the potential for improvement after three years in HE. After having learned IL, several informants expressed feeling more open-minded and more capable of withholding their emotions and prejudices when rendering judgement. Several felt that being information literate was situation-dependent, a competence to exercise and display in an educational environment, but not necessarily in social situations.

These quotes illustrate how students' perceptions of themselves as information-literate people change over time. Quotes are translated from Norwegian, and the names of informants are changed to protect their identities. “Becoming information literate has contributed to my development as a person” (Kari). “IL’s importance has become greater for me. It has an impact on daily life and social life, not just academics” (Magnus). “Being information literate

means a lot to my sense of integrity and to my feeling of being taken seriously when I talk to others” (Heidi).

#### **4.2.2 Thematic analysis**

Based on coding of interview transcriptions, the main research questions, and theoretical frameworks (see Literature review), five themes were created, each representing semantic patterns in the interview data. In alphabetical order, these themes are (a) Changes in identity as an information literate person; (b) Development in IL attitudes; (c) Development in IL doing; (d) Development in IL knowing; and (e) Interest in being or becoming information literate. Themes are presented in detail in Paper 4, with example quotes from interviews. Additional comments and quotes are found below.

##### a) Changes in identity as an information literate person

Most informants considered IL more instrumental – a kind of tool – than something related to themselves as people, i.e., their *identity*. However, when interpreting what they actually said, becoming more information literate has clearly affected them as individuals. They expressed having become critical thinkers and more able to distinguish between objective and subjective sources. They feel that they now have more integrity, both academically, when creating knowledge, and in discussions with peers. They experience that they can better manage and withhold their emotions, allowing them to express their point of view and argue more effectively, and to better recognize and evaluate information sources that evoke emotional responses. Instead of being quick to judge, they now consciously and critically reflect on content before making their assessments. They expressed feeling more competent and confident in certain situations, both at the university and in daily life, regarding their abilities to find, process, and communicate information. They felt proud of themselves as information-literate people.

These developments indicate that some informants have undergone significant changes as individuals, although they may not have used the word *identity* in their descriptions of these changes. When referring to Erikson’s (1968) description of personal identity – our perceptions of ourselves as being the same in all situations and our perceptions of how others judge us – it can be argued that some informants have undergone changes in identity. One example is Amy, who said the following as a result of having become more information literate: “I’m not the same person today as I was before. [Previously, I was] perhaps somewhat

less reflective ... [and became] emotionally engaged a little faster.” Magnus provides another indication of an identity change: “I see it as an advantage to be information literate, and see clearly how it connects to identity and how I relate to different things.” The identity changes that Amy and Magnus describe fall into Illeris’s *personality layer* of identity, where he posits that TL is concentrated (Illeris, 2014b). This layer concerns how one relates to others and to the outside world, and includes personality traits such as behavior, values, and attitudes, where identity can change in response to new impressions.

#### b) Development in IL Attitudes

Several informants expressed that their attitudes about being information literate had changed, especially their sense of IL’s importance and usefulness. Amongst other things, they felt that IL is important when choosing which information sources to rely on, and useful for achieving passing grades or making health-related decisions. For example, Amy told that she had “actually changed some opinions about things” after learning more about source discernment. They had less patience for those who Zurkowski might call *information illiterate*, including friends, partners, and family members.

The experience of becoming a critical thinker – seeking reasons on which to base one’s actions and judgements (Bates, 2019) – was also expressed in interviews. Ingrid told that instead of only using Google, she has now become “more critical to the search tools I use.” Kari said, “I used to be very gullible ... [but] now I have less trust in, and am more critical of, the sources I find.” Participants also spoke about how they can turn on and off their IL, dependent on the situation. They would likely turn on their IL in a colloquium group at the university, but may turn it off in a social situation with non-academics, to not appear overly critical. They recognize that it is not always necessary to agree with others’ opinions – of a misinformed Facebook post, for example – and that it is not always worth an ensuing argument. Their attitudes regarding being information literate may thereby be situation-dependent. Other changes in attitudes, such as becoming more open-minded, curious, and gaining more of a sense of integrity, were also expressed by informants.

#### c) Development in IL Doing

The interview guide included questions regarding students’ perceived development in their abilities to find, evaluate, and use information sources. All informants felt that these skills had developed substantially over three years, and they referred especially to their improved

abilities to find and evaluate sources. Many mentioned that prior to entering the university, they had searched for information mainly in Google and Wikipedia. After being introduced to Google Scholar and the library's databases and discovery system, as well as learning and applying useful search techniques, they can now search more effectively and thus find more reliable and relevant sources of information. One informant, Eli, described learning about the PsycINFO database and search technique as "priceless" for her.

d) Development in IL Knowing

All informants told of considerable growth in their IL knowledge since starting at the university. They felt that their academic competence had increased – they had become better able to produce valid knowledge and to corroborate their arguments. The tasks they were given motivated the development of their IL knowledge and skills, i.e., *doing* fueled interest and learning. IL was often seen as something instrumental in helping them to succeed as a student. Those students who planned on going to graduate school were most interested in continuing their IL learning, since they had use for IL knowledge and skills in their future studies. Students told that the knowledge they acquired in the classroom contributed to their becoming more critical and reflective thinkers, and enabled them to search for, evaluate, and use information sources more effectively. Heidi demonstrated this when she said, "I'm knowledgeable enough that I don't drink bleach to avoid coronavirus when a president says so!" Others also pointed out various risks of *not* being information literate, such as unintentionally plagiarizing, believing in conspiracy theories, refusing to be vaccinated for COVID, or even losing their life in an avalanche (an actual danger in this part of the world).

e) Interest in being or becoming information literate

Rather than asking students about their interest *in IL*, we asked about their interest *in being or becoming information-literate people*, which is relevant to their identity. Most informants expressed more interest in *being* than *becoming* information literate, as they can achieve their goals as a student by being information literate, and they also like the feeling of being information literate. Outside of academia, they did not have this interest. The only informants who expressed an interest in *becoming* information literate, were those who planned to continue their education.

### 4.3 Mixed methods findings

QUAN and QUAL results at times confirmed each other, while at other times, there was a discrepancy between the two. Mixing methods can thereby provide additional insight and reveal more interesting aspects of undergraduates' IL development over time than either method could provide on its own. This is demonstrated in Paper 4, Table 3, where results from both methods are integrated and exemplified by student quotes, and in the paragraphs below.

Although QUAN results from surveys and assignment-based measures show some positive, significant correlations between what students *know* and *do*, QUAL results show that students do not consistently practice their IL knowledge in all situations. When writing a paper, they are more careful about critically evaluating their information sources than when conversing casually with friends or family. As Beth said, “especially when there are a lot of feelings in the picture, there can be differences between what you know and what you do.”

In Paper 3's metacognition study, lower-performing students overestimated their IL abilities, while higher-performing students underestimated them (Dunning-Kruger effect). In interviews, students described a similar phenomenon, namely an initial over-confidence in their IL abilities. Fred illustrated this when he said that, at first “I was a little over-confident in my abilities... Now I'm more unsure than I was before.” Magnus told that “the more you learn, the more you realize all you don't know.”

Using the interest-o-meter (Appendix 9) to depict their interest trajectories, informants indicated that their interest in being or becoming information literate increased with time, although in different ways. When comparing these results TRIQ results for mean interest phase, there is a discrepancy in QUAN and QUAL findings. However, when comparing them with TRIQ results from the Meaningfulness subscale, findings reinforce each other.

## 5 Discussion

Undergraduates are young adults in the process of learning not only a specific discipline, but also how to function in a constantly changing information environment. Students require knowledge, skills, and attitudes to enable them to become responsible producers and consumers of information, earn a living, navigate new demands and challenging situations, become responsible citizens, and continue their paths of lifelong learning. Becoming information literate is an important part of this journey.

### 5.1 Contributions to the field

This dissertation contributes to the body of IL research, and to its practice, in several ways. It provides new tools to measure various aspects of IL and fills gaps in the IL literature regarding the use of mixed methods, and in studying the development of relationships between IL knowledge, skills, and attitudes over time. Additionally, it reestablishes the construct of interest, an important motivator for learning (Kim & Schallert, 2014; Renninger & Hidi, 2016) into the IL literature. The exploration of student learning through the lens of transformative learning theory also fills a gap, as previous research has focused mainly on transformative experiences of those who *teach* IL. These valuable and unique contributions to the field of IL are enumerated below.

First: addressing the need for better tools to measure IL (see e.g., Hollis, 2018; Mahmood, 2017). The first study in this project resulted in the creation of a suite of freely available, psychometrically evaluated tools for assessing IL knowledge, skills, and attitudes in undergraduates – the Tromsø Information Literacy Suite (TROILS). These tools are applicable to most disciplines, either in connection with IL teaching or for research purposes. The 21-item IL-knowledge test TILT can be employed to quickly and easily assess knowledge of source-based aspects of IL (finding, evaluating, and using information). Results show that the act of taking the test has value in itself, especially for students at the start of their undergraduate education, as it helps make them aware of their IL levels (see Paper 3, Figure 1). For IL practitioners, testing will help to identify those students who know the least, and to discover what type of knowledge is lacking. TILT can be repeated on multiple occasions to assess student progress or to evaluate the effect of IL instruction or other interventions. The two assignment-based measures in this research are useful for assessing IL skills in practice and can be utilized with students in any discipline who are writing a research paper. The interest measure TRIQ is created in two versions – one is object-specific, for

measuring interest in being or becoming an information literate individual – and the other is object-general, and can be used to measure interest in any object of interest. For those who study interest, this questionnaire might prove valuable.

Second: addressing the dimensionality of the IL construct. Factor analyses in this project's first study suggest that IL is a broad, multifaceted construct, thereby confirming findings from Beile O'Neil (2005) and Morley (2014). IL's homogeneity is otherwise often implied in the IL literature. IL's heterogeneity has several implications for the field, including the way IL is measured. For example, general tests of IL should be considered indexes rather than scales, and their reliability should therefore *not* be assessed using internal consistency measures. Also, scores on individual items from IL tests may be informative in their own right, indicating specific aspects of IL knowledge, rather than as mere contributors to an overall IL level.

Third: this research reestablishes the construct of interest, an important motivator for learning (Kim & Schallert, 2014; Renninger & Hidi, 2016), into the IL literature. With the exception of this doctoral research and the research of Black (2018), no others – as far as I have found – have explored the connection between IL and interest. Results from the second study demonstrate that the effort students intend to put into learning is more highly associated to their interest than to their perceived learning need. This implies that by developing interest in IL, whether through incorporating IL skills into assignments or by focusing on its relevance when teaching, IL teachers can incite students' motivation to learn. Results from the interest questionnaire TRIQ offer insight into how IL interest develops, thereby providing us with clues as to the kind of support students may need at each phase of their interest development. This is valuable information for IL educators who must first attempt to trigger IL interest in students and later to maintain this interest over time. IL researchers or practitioners can choose which TRIQ subscales they wish to employ, and thereby measure a variety of variables including positive affect, meaningfulness, and learning aspirations.

Fourth: addressing factors that affect students' self-awareness of their IL abilities. This research augments previous IL studies involving metacognition by exploring associations among students' estimated IL abilities with variables such as education level, gender, performance level, and the timing of estimates (before or after taking the IL test). For example, our results show that PhD students have better metacognitive awareness of their

abilities than undergraduates, that men have higher and more accurate IL test score estimates than women, and that the act of taking the test increases metacognitive awareness.

Fifth: addressing the growth and integration of IL knowledge, skills, and attitudes over time. This research is the first, as far as I can see, that explores interactions between *knowing*, *doing*, and *feeling* over several years. In the third study's cross-sectional sample of undergraduate psychology students, significant increases in IL knowledge (*knowing*) and skills (*doing*) were found between their first and sixth semester. Significant growth in *feeling*, when represented by TRIQ's Meaningfulness subscale, was also observed. Correlations between *knowing*, *doing*, and *feeling* increase over three years as well, but not significantly.

Sixth: addressing transformative learning in students as a result of becoming information literate. This doctoral research explores TL in *students*, whereas previous IL research has focused mainly on TL experiences among those who *teach* IL. IL is valuable not only instrumentally – to do a better job – but also existentially – to become a more reflective and self-aware person. Becoming information literate may in this way be one of the central factors in HE that contributes to students' abilities to manage the challenges of daily life in the information age. If one of the aims of HE is this type of holistic development in students, knowing whether TL takes place is important.

Seventh: addressing the value of mixed methods in IL research. The previous point – where inferences were made by integrating the QUAN and QUAL evidence – illustrates how the use of mixed methods can enhance IL research. This methodology contributes to a more comprehensive understanding of IL development than possible with QUAN or QUAL methods on their own. The application of mixed methods in IL research, however, is limited.

## **5.2 Development over time**

This was not an experimental study, but rather an explorative study to measure change over time without a specific intervention. Results from *know* and *do* measures show significant growth in IL knowledge and skills in undergraduates over a three-year period. Although some readers may consider this an obvious conclusion and not necessary to measure or report, I argue that this growth is *not* inevitable, and that it should be documented in valid and reliable ways. Not all students have learned IL skills in their undergraduate education, as the emphasis on IL instruction may vary between institutions, faculties, and disciplines, and also because some students put more effort into learning IL than others. IL growth should therefore not be



expected, but must be observed and documented by employing psychometrically evaluated tools such as those developed in this project.

As seen in Section 4.2-Metacognition, results indicate that students' metacognitive abilities decreased over time. Two factors may help to explain this result. First, there was a great degree of variability in the data, as indicated by large standard deviations. Second, the analyses over time are based on *after*-test score estimates since students were not asked to make *before*-test estimates in their first semester. Because this result was unexpected, it would be valuable to replicate this study in the future. To increase the value of this replication study, *before*-test score predictions and questions related to interest/need/effort should be included.

Findings show that students' knowledge (*know*) and skills (*do*) increase significantly – two aspects of their IL trajectory based on measures that tap concrete evidence of growth. Student interest (*feel*), however – the most subjective of our measures – shows more varied results. Over the course of three years, achieving sufficient IL levels to satisfy academic goals may quench student interest in becoming more information literate. Accordingly, this interest may wane. However, achieving such levels may nevertheless inspire other students to maintain, or even increase, their interest, hence the varied results. The observed increases over time in TRIQ subscales for Competence level and Meaningfulness make sense intuitively, since many students become more proficient with their IL skills and more aware of IL's importance and usefulness with time. However, not all TRIQ subscale results were in accordance with each other, either intuitively or as described by theory. For example, TRIQ results show that as Meaningfulness subscale scores increased, Competence aspiration subscale scores decreased. One would intuitively expect that as an interest becomes more meaningful, aspirations associated with that interest likewise would grow, and this is also predicted by the four-phase model of interest (Hidi & Renninger, 2006). In our case however, results can perhaps be attributed to students being more aspiring at the start – when they still had a lot to learn – than after having learned much of what they needed to succeed as a student.

Divergence was also found between QUAL and QUAN results. In interviews, nearly all informants described an increase in their interest in being or becoming information-literate people over three years, as did those few who plotted their interest with the interest-o-meter. Some TRIQ subscale scores, including the score for Meaningfulness, corroborate the growth they describe, however, results from other TRIQ subscales, such as Competence aspiration,

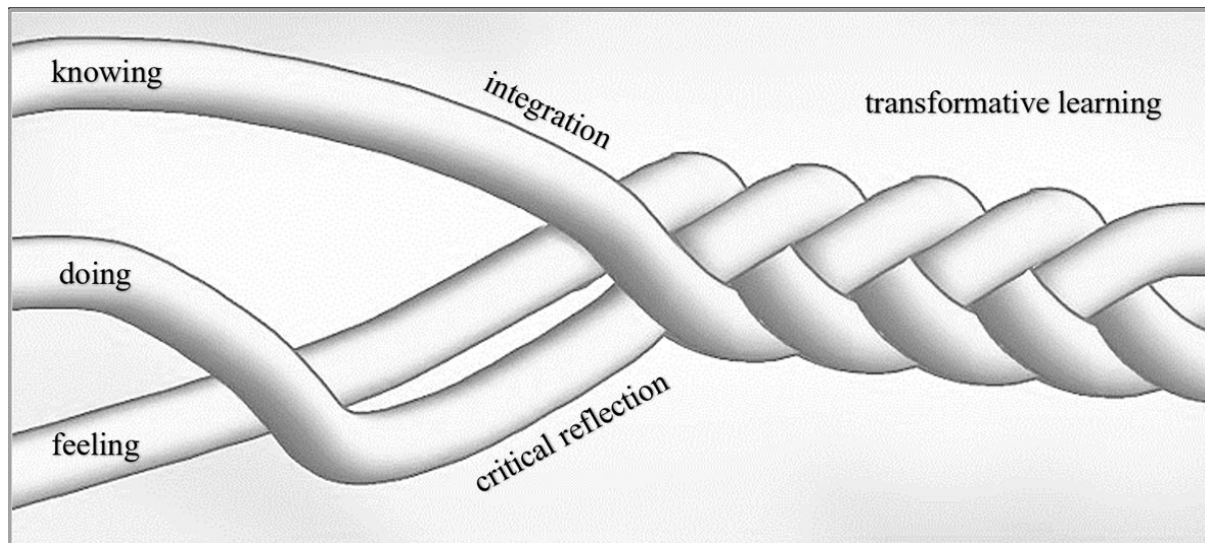
are inconsistent with the expressed growth. Diverging results such as these are common in mixed-methods research and are sometimes welcomed by the researcher as an analytical challenge that may bring forth insight (Bazeley, 2018, p. 264). Possible explanations for the discrepancy between QUAN and QUAL results are that interview volunteers may have been more interested in IL than others, or that they exaggerated their interest in the interviews in order to meet the interviewer's (imagined) expectations. As stated above, another explanation for the concurrent increase in interest and decrease in aspirations, may be that when people feel they have reached a desired level of competence, their need to aspire to learn more might naturally abate.

### **5.3 Integration and transformative learning**

In addition to measuring the development of IL *knowing*, *doing*, and *feeling* separately, valuable information regarding students' development as information-literate individuals is gained by examining associations among the three. QUAN results provide evidence that IL *knowing*, *doing*, and *feeling* become more strongly integrated, and less fragmented, as students progress in their education. This is illustrated by circle overlaps in Figure 7, showing the increase in correlations over time between scores on TILT, *do* measures, and Meaningfulness. This overlap is where I imagine TL taking place, where knowledge, skills, and attitudes become an integrated whole. This can affect our personalities – how we think, behave, and feel – and lead to transformation. Mezirow would consider this integration an essential element of *transformative learning*, which he wrote includes an “...integrative understanding of one's experience” (Mezirow, 1990, p. xvi). Baartman and de Bruijn (2011) would call it *transformative integration*, a type of learning that leads to an enhancement of competence and understanding. An example of this integration is when students, in assignments in which IL skills are assessed, apply their IL knowledge and are influenced by IL attitudes such as their understandings of the ethical issues involved in use of information.

I envision the integration of knowledge, skills, and attitudes (*knowing*, *doing*, and *feeling*) over time as a braiding of IL learning (see Figure 8), that may in some students result in a transformative learning experience. The critical reflection of the knowledge, skills, and attitudes involved in learning is an important ingredient in this transformation.

Figure 8: Transformative Learning: An Integration of Knowing, Doing, and Feeling, Involving Critical Reflection



Note: Adapted from *How to Braid Rope*, of Nicole Bolin, 2021, wikiHow (<https://www.wikihow.com/Braid-Rope>). CC BY-NC-SA 3.0

In this figure, *feeling* consists of both the QUAN and the QUAL data, since by themselves, neither provides adequate evidence of a TL experience in students. However, when integrating the QUAN and QUAL results, I believe that there is a sufficient indication of a change in identity – which implies TL, according to Illeris (2014b) – in several students. In interviews, several students expressed a change in identity to some extent as a result of their becoming more information literate. Some spoke explicitly of an identity change, while others expressed an increased ability to think critically and to reflect on their previous assumptions and beliefs surrounding information. This is illustrated by a quote from Ingrid, who speaks of critical self-reflection, integrity, the critical evaluation of sources, and ethical issues involved in the dissemination of information:

Since the start of my studies, I've gradually become more reflective about the things I do. The more assignments I have, the more I realize how important it is that I do them in the correct way and that I think about the sources I've used... I wasn't as critical of sources before starting at the university. I [now] think more about the intentions of people who write articles, if they're trying to get some [hidden] message across, like that the Earth is flat.

Illeris and Mezirow may consider Ingrid to be on a trajectory toward TL, as she demonstrates self-reflection, critical thinking, and a perspective transformation leading to changes in attitudes and behavior. Her development is taking place in what Illeris would consider the personality layer of identity, as her behavior, values, and attitudes are changing in response to new impressions. Evidence strengthening the claim that TL has occurred is provided in Paper 4 (Discussion – Research questions). Characteristics of TL that I see evidence of in some student interviews include that TL refers to learning: in adults; after critical self-reflection; that leads to profound changes in perspectives and behavior; that leads to changes in identity; that is irreversible; and that implies a qualitative change in the learner, beyond the acquisition of new knowledge and skills.

This quote shows how the use of QUAL methods in this project has augmented the QUAN findings. Mixed methods has thereby contributed to answering research questions about the development of IL and to discovering significant learning – possibly transformative – in students as a result of their becoming increasingly information literate.

## **5.4 Relevance for IL instruction**

Ingrid's quote is an encouraging reminder to those of us who teach IL that our work is important. It shows how IL *feeling* influences learning, and it underscores the importance of emphasizing the *why's* of IL (e.g., why we cite sources), and not only the *how's* (e.g., use of particular reference style), in IL instruction. I believe that our instruction could improve by becoming familiar with theories of TL. If the aim of our teaching is for students to internalize the values involved in IL, for example being critical of information or duly crediting the work of others, then some students who internalize these values may experience a type of transformation. Being aware of this aim of internalization, and of its possible transformative effects in students, can make us better teachers and more likely to achieve this objective, as we can then design teaching, course work, and assessment based on this goal. In addition, our teaching may benefit by taking into consideration IL's transferability, how it applies to tasks also outside of HE.

Interest is important for sustaining engagement in a topic and motivating learning (Hidi & Renninger, 2006; Kim & Schallert, 2014). Students' interest in IL can be triggered if they realize that being information literate not only helps them academically, but is also useful in their personal life beyond the university. In this research, we found that students' interest in being or becoming information literate can develop early in their studies, and that interest

levels can wax, wane, or remain stable over time. It is therefore important for IL instructors who meet first-year students to explicitly take into account their interest, in addition to the specific IL knowledge and skills to be mastered.

As students told of in interviews, their IL learning was bolstered by putting their newly learned IL skills into practice, especially when these skills were assessed in coursework. In other words, *doing* fuels learning. According to Secker and Coonan (2011) and Løkse et al. (2017), this integration of theory and practice is best achieved when librarians and other teaching faculty collaborate in formulating IL-relevant learning goals in curricula, in designing assignments that require IL skills, and in assessing the achievement of learning goals, formatively and consistently, in coursework throughout a study program. Many IL researchers (e.g., Løkse et al., 2017; Secker & Coonan, 2011) maintain that this type of constructive alignment (see Biggs, 2014) is best achieved with embedded IL instruction – where IL is integrated into course content, learning activities, and assessment – rather than with conventional one-shot instruction.

IL should be considered an important component of HE by all who design curricula, teach, and assess students. IL is relevant for nearly all disciplines in the humanities, natural sciences, and social sciences – in other words, IL is *pandisciplinary*. It is partially because IL is important for all students, that academic libraries – who serve the entire institution – often take on the responsibility of teaching IL in higher education.

## **5.5 Reflections on research process**

Part of doing research is reflecting upon the process as a whole. What advice would I give to others starting longitudinal research? What could have been done to facilitate the research process? How could the research have been enhanced? Did I achieve my research goals?

First – regarding advice to others planning longitudinal research – avoid pandemics! In addition to affecting students' mode of instruction (digital vs. physical) and physical and emotional health, I have reason to believe that the COVID pandemic also contributed to reduced participation in surveys and perhaps to students aborting or delaying their educational trajectories.

Second, two modifications may have facilitated the research process.

- a) It would have been beneficial to collect cross-sectional data from parallel cohorts concurrently with longitudinal data from the start of the project as a back-up plan. This would assure larger sample sizes and possibly a better statistical basis for conclusions and inferences, but it would mean that the research is not purely longitudinal.
- b) Instead of sending surveys only to those students who had responded to previous surveys – as I did at the end of the first and fourth semesters – it would have been advantageous to send each survey to all students who started their studies fall 2019. I could thereby have procured additional data from the longitudinal group, although this data may have been incomplete.

Third, the research may have been enhanced by assessing the source-evaluation assignment in another way, rather than by using a combination of quality, frequency, and variety subscores. Total source-evaluation scores on this assignment did not correlate significantly with TILT scores, indicating that it may not sufficiently discriminate among students with different IL-knowledge levels.

Fourth, the research goals established at the beginning of the PhD project were ambitious and only partially achieved. Aims enumerated in the original research protocol included (1) developing and validating measures for IL knowledge (*know*), skills (*do*), and interest (*feel*); (2) determining how a cohort of students who had embedded IL instruction in their first semester developed over three years, in light of their *knowing*, *doing*, and *feeling*; and (3) comparing IL development in students with embedded IL instruction to those with other types of instruction. The first two of these aims have been achieved, while the third – which will help us to make more informed decisions about how to best teach IL – remains to be done. Based on the research performed in this PhD project, we now have tools and detailed knowledge that help us to better understand what students *know*, *do*, and *feel*. These, in turn, will allow us to better explore the most effective ways to teach IL.

## 5.6 Limitations and considerations

Limitations of the individual studies in this project are addressed in the four papers. When addressing the project as a whole, there are some general limitations and various factors to consider. One is the small sample size in the longitudinal repeated-measure study, which makes certain analyses, such as regressions, impossible to perform with a sufficient degree of statistical power. Students came and went over the course of the project. Fewer than 30% of

the students in the original 2019-cohort were enrolled in the third year of the bachelor's or professional-studies program three years later,<sup>11</sup> and not all of these participated in surveys. Decreased sample size over time, a common feature of longitudinal studies, may in this case have been amplified by the pandemic. COVID restrictions were effectuated only six months after this project's longitudinal cohort began their undergraduate studies, and their classes were online for nearly two years.

Another factor that may have influenced outcomes in this project is that most of its studies included psychology students as respondents. These students may have more insight into concepts such as identity or interest than students in other disciplines. In addition, the psychology department at UiT emphasizes reference technique and the correct use of sources to a greater extent than many other departments. These factors may make our research findings important to test with other student samples from various disciplines, and they prevent us from generalizing our findings for the entire population of undergraduates.

It is worth reminding the reader of a feature of the instruments developed in this project for measuring IL *knowing, doing, and feeling*. We have psychometrically evaluated these instruments and made them freely available<sup>12</sup>, and we hope that they will be used by other IL researchers and practitioners. However, since the measures' psychometric properties have been evaluated in a specific situation, they should be further tested for general use, as evidence of validity and reliability should ideally be established for each specific sample and circumstance (Streiner & Kottner, 2014).

Although this is not a limitation, it is also worth mentioning the possibility of familywise error in any *family of tests*, i.e., tests conducted on the same set of data, addressing the same research questions. Type I error can occur when a researcher mistakenly rejects the null hypothesis and falsely concludes that there is an effect in a sample when one does not exist in a population, i.e., a *false positive* finding. The probability of making a Type I error increases in studies containing a family of tests (Field, 2018, p. 83). Decisions made along the way can

---

<sup>11</sup> Exact numbers were not possible to obtain. Several students in the original 2019-cohort had planned to study psychology for only one year, and did not continue in the bachelor's or professional-studies program.

<sup>12</sup> See <https://site.uit.no/troils>

compound each other, and thereby increase the probability that chance patterns may be found. This type of error is then referred to as *familywise error*, another factor that can make the replication of some studies – especially those with multiple analyses and inferences – more problematic. The study in Paper 4 includes many comparisons and statistical inferences, and I have attempted to avoid familywise error by carefully considering several factors, including which data to use, how to analyze them statistically, and which inferences can be made based on the evidence. An example of where familywise error may have played a role is in my inference that TL has occurred in some students. TL is difficult to measure and ascertain, and inferences of its existence depend on several choices, including which definition of TL to base the study on, which methods to employ, which questions to pose, which data to include, which analyses to perform, and how results are integrated and interpreted.

## **5.7 Future research**

This project has spurred several ideas for future research, including comparing IL development over time in (1) undergraduates with embedded versus one-shot IL instruction in their first year; (2) undergraduates with IL instruction in the first versus second semester; and (3) undergraduates with IL instruction every year versus less frequently. It would also be fruitful to (4) replicate the metacognition study to examine the development of students' IL metacognitive abilities over time, as some of our results were unexpected; (5) compare students' IL levels with other benchmarks of student achievement such as grades or rates of completion, and to (6) perform a case study following individual students over a longer time period, e.g. 5 years, to explore how their IL knowledge, skills, and attitudes continue to develop after their undergraduate education, and how their IL competence takes them beyond the context in which it was learned. The tools for measuring IL developed in this doctoral research can be employed in these future studies, and mixed methods can be used in some cases to better illuminate the development of IL attitudes and social and emotional components of IL learning. An idea for future QUAL research is to (7) focus on the ethics of IL – how ethical issues are addressed in IL teaching and the importance of ethical considerations for students.



## 6 Conclusion

This article-based dissertation entails both research and development in information literacy. Valuable tools for measuring IL knowledge, skills, and attitudes were developed, including an IL-knowledge test, two assignment-based measures for assessing the evaluation and use of sources, and an interest questionnaire. These tools are psychometrically tested for reliability and validity and are freely available for use by IL practitioners and researchers.

The dissertation sheds light on several knowledge gaps in the IL literature, including (1) the development of IL in students over a period of several years; (2) changes in the associations among IL knowledge, skills, and attitudes over time; (3) connections between IL and interest; (4) factors affecting students' self-awareness of their IL abilities; and (5) transformative learning in students as a result of their becoming more information literate. The use of mixed methods in this dissertation also represents a valuable contribution to the field.

A final quote from an interview conveys the importance of IL in an undergraduate education and demonstrates why it is a worthy topic of research, because as Eli says, "IL underpins everything else I do academically."

## References

- Akkerman, S. F., & Bakker, A. (2019). Persons pursuing multiple objects of interest in multiple contexts. *European Journal of Psychology of Education*, 34(1), 1-24. <https://doi.org/10.1007/s10212-018-0400-2>
- American Educational Research Association. (1999). *Standards for educational and psychological testing*. American Educational Research Association.
- American Library Association. (1989). *Presidential committee on information literacy: Final report*. <http://www.ala.org/acrl/publications/whitepapers/presidential>
- Association of College and Research Libraries. (2000). *Information literacy competency standards for higher education*. <http://hdl.handle.net/11213/7668>
- Association of College and Research Libraries. (2015). *Framework for information literacy for higher education*. <https://www.ala.org/acrl/standards/ilframework>
- Bates, B. (2019). *Learning theories simplified ... and how to apply them to teaching* (2nd ed.). SAGE.
- Bauman, Z. (2000). *Liquid modernity*. Polity Press.
- Bazeley, P. (2018). *Integrating analyses in mixed methods research*. SAGE. <https://doi.org/10.4135/9781526417190>
- Beile O'Neil, P. (2005). *Development and validation of the Beile Test of Information Literacy for Education (b-tiled)*. University of Central Florida. <http://purl.fcla.edu/fcla/etd/CFE0000749>
- Berg, C. B., Charles, L. C., Chudnick, S., Cook, H., Dalal, H., Dempsey, M., ... Ghajar, M. (2015). *An open letter regarding the framework for information literacy for higher education*. Retrieved from <https://acrlog.org/2015/01/07/an-open-letter-regarding-the-framework-for-information-literacy-for-higher-education/>
- Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research1. In A. T. Tashakkori, Charles (Ed.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 95-118). SAGE. <https://doi.org/10.4135/9781506335193.n4>
- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1(1), 5-22.

- Black, S. (2018). Development, interest, self-direction and the teaching of information literacy [Article]. *Communications in Information Literacy*, 12(2), 203-214. <https://doi.org/10.15760/comminfolit.2018.12.2.9>
- Boruff, J. T., & Harrison, P. (2018). Assessment of knowledge and skills in information literacy instruction for rehabilitation sciences students: A scoping review. *Journal of the Medical Library Association*, 106(1), 15-37. <https://doi.org/10.5195/jmla.2018.227>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbooks in psychology: APA handbook of research methods in psychology* (Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological, pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). SAGE.
- Baartman, L. K. J., & de Bruijn, E. (2011). Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence. *Educational Research Review*, 6(2), 125-134. <https://doi.org/10.1016/j.edurev.2011.03.001>
- Calfee, R. (1993). Assessment, testing, measurement: What's the difference? *Educational Assessment*, 1(1), 1-7. [https://doi.org/10.1207/s15326977ea0101\\_1](https://doi.org/10.1207/s15326977ea0101_1)
- Christie, M., Carey, M., Robertson, A., & Grainger, P. (2015). Putting transformative learning theory into practice. *Australian Journal of Adult Learning*, 55(1), 9-30.
- CILIP Information literacy group. (2018). *CILIP Definition of Information Literacy 2018*. <https://www.cilip.org.uk/news/421972/What-is-information-literacy.htm>
- Clarke, K. E., & Radcliff, C. (2018). Information literacy assessment for instruction and improvement and demonstration of library value: Comparing locally-grown and commercially created tests. <https://commons.lib.jmu.edu/letfspubs/161/>
- Dahl, T. I., & Nierenberg, E. (2021). Here's the TRIQ: The Tromsø Interest Questionnaire based on the four-phase model of interest development. *Frontiers in Education*, 6(402), 1-17. <https://doi.org/10.3389/feduc.2021.716543>
- Downey, A. L. (2014). *The state of the field of critical information literacy in higher education* (Publication Number 10032249) [Ph.D., University of North Texas]. ProQuest Dissertations & Theses Global. Ann Arbor.

- <https://www.proquest.com/dissertations-theses/state-field-critical-information-literacy-higher/docview/1774248567/se-2?accountid=17260>
- Doyle, M., Foster, B., & Yukhymenko-Lescroart, M. A. (2019). Initial development of the Perception of Information Literacy Scale (PILS). *Communications in Information Literacy*, 13(2), 205-227. <https://doi.org/10.15760/comminfolit.2019.13.2.5>
- Erikson, E. H. (1968). *Identity: Youth and crisis*. Faber & Faber.
- Erlinger, A. (2018). Outcomes assessment in undergraduate information literacy instruction: A systematic review. *College & Research Libraries*, 79(4), 442. <https://doi.org/https://doi.org/10.5860/crl.79.4.442>
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Health Services Research*, 48, 2134-2156. <https://doi.org/10.1111/1475-6773.12117>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
- Freire, P. (1970). *Pedagogy of the oppressed*. Seabury Press.
- Gawlik, K. (2018). Focus group interviews. In M. Ciesielska & D. Jemielniak (Eds.), *Qualitative methodologies in organization studies: Methods and possibilities* (Vol. II, pp. 97-126). Springer International Publishing. [https://doi.org/10.1007/978-3-319-65442-3\\_5](https://doi.org/10.1007/978-3-319-65442-3_5)
- Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Polity Press.
- Granikov, V., Hong, Q. N., Crist, E., & Pluye, P. (2020). Mixed methods research in library and information science: A methodological review. *Library & Information Science Research*, 42(1), 1-20. <https://doi.org/10.1016/j.lisr.2020.101003>
- Gross, M., & Latham, D. (2012). What's skill got to do with it? Information literacy skills and self-views of ability among first-year college students. *Journal of the American Society for Information Science and Technology*, 63(3), 574-583. <https://doi.org/10.1002/asi.21681>
- Gross, M., Latham, D., & Julien, H. (2018). What the framework means to me: Attitudes of academic librarians toward the ACRL framework for information literacy for higher education. *Library & Information Science Research*, 40(3), 262-268. <https://doi.org/https://doi.org/10.1016/j.lisr.2018.09.008>
- Hesse-Biber, S. N., Rodriguez, D., & Frost, N. A. (2015). A qualitatively driven approach to multimethod and mixed methods research. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford handbook of multimethod and mixed methods research inquiry* (pp.

- 3-20). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199933624.013.3>
- Hetland, A., & Vittersø, J. (2012). The feelings of extreme risk: Exploring emotional quality and variability in skydiving and BASE jumping. *Journal of Sport Behavior*, 35(2), 154-180.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111-127.  
[https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4)
- Hoggan, C. D. (2016). Transformative learning as a metatheory: Definition, criteria, and typology. *Adult Education Quarterly*, 66(1), 57-75.  
<https://doi.org/10.1177/0741713615611216>
- Hollis, H. (2018). Information literacy as a measurable construct: A need for more freely available, validated and wide ranging instruments. *Journal of Information Literacy*, 12(2). <https://doi.org/https://doi.org/10.11645/12.2.2409>
- Hooper, M. D. W., & Scharf, E. (2017). Connecting and reflecting: Transformative learning in academic libraries. *Journal of Transformative Education*, 15(1), 79-94.  
<https://doi.org/10.1177/1541344616670033>
- Illeris, K. (2004). Transformative learning in the perspective of a comprehensive learning theory. *Journal of Transformative Education*, 2(2), 79-89.  
<https://doi.org/10.1177/1541344603262315>
- Illeris, K. (2014a). *Transformative learning and identity*. Routledge.
- Illeris, K. (2014b). Transformative learning and identity. *Journal of Transformative Education*, 12(2), 148-163. <https://doi.org/10.1177/1541344614548423>
- Illeris, K. (2017). Transformative learning as change and development of identity. In A. Laros, T. Fuhr, & E. W. Taylor (Eds.), *Transformative learning meets bildung* (pp. 179-190). SensePublishers. [https://doi.org/10.1007/978-94-6300-797-9\\_15](https://doi.org/10.1007/978-94-6300-797-9_15)
- Irwing, P., Booth, T., & Hughes, D. J. (2018). *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development*. John Wiley & Sons.
- Ivanitskaya, L., Laus, R., & Casey, A. M. (2004). Research Readiness Self-Assessment: Assessing students' research skills and attitudes. *Journal of Library Administration*, 41(1-2), 167-183. [https://doi.org/10.1300/J111v41n01\\_13](https://doi.org/10.1300/J111v41n01_13)

- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.  
<https://doi.org/10.3102/0013189X033007014>
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.  
<https://doi.org/10.1177/1558689806298224>
- Julien, H., Gross, M., & Latham, D. (2018). Survey of information literacy instructional practices in US academic libraries. *College & Research Libraries*, 79(2), 179.  
<https://doi.org/https://doi.org/10.5860/crl.79.2.179>
- Kegan, R. (2000). What "form" transforms? A constructive-developmental approach to transformative learning. In J. Mezirow (Ed.), *Learning as transformation: Critical perspectives on a theory in progress* (pp. 35-70). Jossey-Bass.
- Kenney, B. (2008). *The transformative library: A narrative inquiry into the outcomes of information use* (Publication Number 3376039) [Ph.D., University of North Texas]. ProQuest Dissertations & Theses Global. Ann Arbor.
- Kim, T., & Schallert, D. L. (2014). Mediating effects of teacher enthusiasm and peer enthusiasm on students' interest in the college classroom. *Contemporary Educational Psychology*, 39(2), 134-144. <https://doi.org/10.1016/j.cedpsych.2014.03.002>
- King, K. P. (2009). *The handbook of the evolving research of transformative learning: Based on the Learning Activities Survey* (10th anniversary ed.). Information Age Publishers.
- Knowles, M. S. (1988). *The modern practice of adult education: From pedagogy to andragogy* (Rev. ed.). Cambridge Adult Education.
- Kos, D., & Špiranec, S. (2014). Debating transformative approaches to information literacy education: A critical look at the transformative learning theory. In *Information literacy: Lifelong learning and digital citizenship in the 21st century* (pp. 427-435). Springer. [https://doi.org/10.1007/978-3-319-14136-7\\_45](https://doi.org/10.1007/978-3-319-14136-7_45)
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Kvale, S. (2007). *Doing interviews*. Sage.
- Løkse, M., Låg, T., Solberg, M., Andreassen, H. N., & Stenersen, M. (2017). *Teaching information literacy in higher education: Effective teaching and active learning*. Chandos Publishing.

- Mahmood, K. (2017). A systematic review of evidence on psychometric properties of information literacy tests. *Library Review*, 66(6/7), 442-455.  
<https://doi.org/10.1108/LR-02-2017-0015>
- McEvoy, P., & Richards, D. (2006). A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing*, 11(1), 66-78.  
<https://doi.org/10.1177/1744987106060192>
- Meyer, J. H. F., Land, R., & Baillie, C. (Eds.). (2010). *Threshold concepts and transformational learning*. Sense Publishers.
- Mezirow, J. (1978). Perspective transformation. *Adult Education*, 28(2), 100-110.  
<https://doi.org/10.1177/074171367802800202>
- Mezirow, J. (1990). *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning*. Jossey-Bass.
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.
- Mezirow, J. (1994). Understanding transformation theory. *Adult Education Quarterly*, 44(4), 222-232. <https://doi.org/10.1177/074171369404400403>
- Mezirow, J. (Ed.). (2000). *Learning as transformation: Critical perspectives on a theory in progress*. Jossey-Bass.
- Mezirow, J. (2003). Transformative learning as discourse. *Journal of Transformative Education*, 1(1), 58-63. <https://doi.org/10.1177/1541344603252172>
- Morley, S. K. (2014). *Initial development of a medical information literacy questionnaire (Dissertation)* [The University of New Mexico College of Education].  
[https://digitalrepository.unm.edu/educ\\_ifce\\_etds/30](https://digitalrepository.unm.edu/educ_ifce_etds/30)
- Nichols Hess, A. (2018). *Transforming academic library instruction: Shifting teaching practices to reflect changed perspectives*. Rowman & Littlefield.
- Nichols Hess, A. (2019). Academic librarians' educational factors and perceptions of teaching transformation: An exploratory examination. *Evidence Based Library and Information Practice*, 14(3), 52-76. <https://doi.org/10.18438/eblip29526>
- Nichols Hess, A. (2020). Academic librarians' teaching identities and work experiences: Exploring relationships to support perspective transformation in information literacy instruction. *Journal of Library Administration*, 60(4), 331-353.  
<https://doi.org/10.1080/01930826.2020.1721939>
- Nierenberg, E., & Fjeldbu, Ø. G. (2015). How much do first-year undergraduate students in Norway know about information literacy? *Journal of Information Literacy*, 9(1), 15-33. <https://doi.org/10.11645/9.1.1983>

- Nierenberg, E., & Dahl, T. I. (2021). Is information literacy ability, and metacognition of that ability, related to interest, gender or education level? A cross-sectional study of higher education students. *Journal of Librarianship and Information Science*, 1-13.  
<https://doi.org/10.1177/09610006211058907>
- Nierenberg, E., Låg, T., & Dahl, T. I. (2021). Knowing and doing: The development of information literacy measures to assess knowledge and practice. *Journal of Information Literacy*, 15(2), 78–123. <https://doi.org/10.11645/15.2.2795>
- Oakleaf, M., Millet, M. S., & Kraus, L. (2011). All together now: Getting faculty, administrators, and staff engaged in information literacy assessment. *portal: Libraries and the Academy*, 11(3), 831-852. <https://doi.org/10.1353/pla.2011.0035>
- Onwuegbuzie, A. J., Johnson, R. B., & Collins, K. M. T. (2009). Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. *International journal of multiple research approaches*, 3(2), 114-139.  
<https://doi.org/10.5172/mra.3.2.114>
- Pearce, L. D. (2015). Thinking outside the Q boxes: Further motivating a mixed research perspective. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford handbook of multi and mixed methods research inquiry* (pp. 42-56). Oxford University Press.  
<https://doi.org/10.1093/oxfordhb/9780199933624.013.4>
- Pinto, M., & Fernández-Pascual, R. (2019). How a cycle of information literacy assessment and instruction stimulates attitudes and motivations of LIS students: A competency-based case study. *Journal of Librarianship and Information Science*, 51(2), 370-386.  
<https://doi.org/10.1177/0961000617742447>
- Renninger, K. A. (2009). Interest and identity development in instruction: An inductive model. *Educational Psychologist*, 44(2), 105-118.  
<https://doi.org/10.1080/00461520902832392>
- Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement*. Routledge.
- Sample, A. (2020). Historical development of definitions of information literacy: A literature review of selected resources. *The Journal of Academic Librarianship*, 46(2), 102116.  
<https://doi.org/10.1016/j.acalib.2020.102116>
- Schilling, K., & Applegate, R. (2007). *Evaluating library instruction: Measures for assessing educational quality and impact*. Paper presented at The Thirteenth National Conference of the Association of College and Research Libraries, Baltimore, MD.  
<http://www.ala.org/acrl/conferences/confsandpreconfs/national/baltimore/baltimore>



- Schilling, K., & Applegate, R. (2012). Best methods for evaluating educational impact: A comparison of the efficacy of commonly used measures of library instruction. *Journal of the Medical Library Association*, 100(4), 258-269. <https://doi.org/10.3163/1536-5050.100.4.007>
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(S2), 107-131. <https://doi.org/10.1007/s11577-017-0454-1>
- Secker, J., & Coonan, E. (2011). *A New Curriculum for Information Literacy (ANCIL) - Curriculum and supporting documents*. <http://www.dspace.cam.ac.uk/handle/1810/244638>
- Silverman, C. (2016). *This analysis shows how viral fake election news stories outperformed real news on Facebook*. BuzzFeed News. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook#.pnEqX1D5R>
- Sobel, K., & Sugimoto, C. R. (2012). Assessment of learning during library instruction: Practices, prevalence, and preparation. *The Journal of Academic Librarianship*, 38(4), 191-204. <https://doi.org/10.1016/j.acalib.2012.04.004>
- Streiner, D. L., & Kottner, J. (2014). Recommendations for reporting the results of studies of instrument and scale development and testing [Special paper]. *Journal of Advanced Nursing*, 70(9), 1970-1979. <https://doi.org/https://doi.org/10.1111/jan.12402>
- Tashakkori, A., Johnson, B., & Teddlie, C. (2020). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences* (2nd ed.). SAGE.
- Tashakkori, A., & Teddlie, C. (2016). *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). SAGE.
- University and University Colleges Act, LOV-2005-04-01-15 Norway (2005). <https://lovdata.no/dokument/NLE/lov/2005-04-01-15>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Wang, V. C. X. (2017). *Theory and practice of adult and higher education*. Information Age Publishing.
- Whitworth, A. (2009). *Information obesity*. Chandos Publishing.

Whitworth, A. (2012). The reflective information literacy educator. *Nordic Journal of Information Literacy in Higher Education*, 4(1), 38-55.

<https://doi.org/10.15845/noril.v4i1.151>

Zurkowski, P. G. (1974). *The Information service environment relationships and priorities: Related paper no. 5*. Washington, DC: National Commission on Libraries and Information Science. Retrieved from <https://files.eric.ed.gov/fulltext/ED100391.pdf>

# Paper 1

# Journal of Information Literacy

ISSN 1750-5968

Volume 15 Issue 2

June 2021

## Article

Nierenberg, E., Låg, T., Dahl, T. I. 2021. Knowing and doing: The development of information literacy measures to assess knowledge and practice. *Journal of Information Literacy*, 15(2), pp.78–123.

<http://dx.doi.org/10.11645/15.2.2795>



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Copyright for the article content resides with the authors, and copyright for the publication layout resides with the Chartered Institute of Library and Information Professionals, Information Literacy Group. These Copyright holders have agreed that this article should be available on Open Access and licensed under a Creative Commons Attribution ShareAlike licence.

"By 'open access' to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."

Chan, L. et al. 2002. Budapest Open Access Initiative. New York: Open Society Institute. Available at: <http://www.soros.org/openaccess/read.shtml> [Accessed: 18 November 2015]

# Knowing and doing: The development of information literacy measures to assess knowledge and practice

Ellen Nierenberg, Research Fellow in Information Literacy, UiT The Arctic University of Norway. Email: [ellen.nierenberg@uit.no](mailto:ellen.nierenberg@uit.no) ORCID: [0000-0001-8666-8092](https://orcid.org/0000-0001-8666-8092)

Torstein Låg, Senior Academic Librarian, UiT The Arctic University of Norway. Email: [torstein.lag@uit.no](mailto:torstein.lag@uit.no) ORCID: [0000-0002-1325-5235](https://orcid.org/0000-0002-1325-5235)

Tove Irene Dahl, Professor of Psychology, UiT The Arctic University of Norway. Email: [tove.dahl@uit.no](mailto:tove.dahl@uit.no) ORCID: [0000-0002-8036-8627](https://orcid.org/0000-0002-8036-8627)

## Abstract

This study touches upon three major themes in the field of information literacy (IL): the assessment of IL, the association between IL knowledge and skills, and the dimensionality of the IL construct. Three quantitative measures were developed and tested with several samples of university students to assess knowledge and skills for core facets of IL. These measures are freely available, applicable across disciplines, and easy to administer. Results indicate they are likely to be reliable and support valid interpretations. By measuring both knowledge and practice, the tools indicated low to moderate correlations between what students *know* about IL, and what they actually *do* when evaluating and using sources in authentic, graded assignments. The study is unique in using actual coursework to compare knowing and doing regarding students' evaluation and use of sources. It provides one of the most thorough documentations of the development and testing of IL assessment measures to date. Results also urge us to ask whether the source-focused components of IL – information seeking, source evaluation and source use – can be considered unidimensional constructs or sets of disparate and more loosely related components, and findings support their heterogeneity.

## Keywords

higher education; information literacy; knowing and doing; Norway; quantitative assessment

---

## 1. Introduction

Information literacy (hereafter IL; roughly, the complex abilities needed to find, evaluate and purposefully use information) is claimed to be important for learning (Oakleaf, 2014), for our empowerment as citizens (CILIP IL group, 2018), for reducing inequality and increasing tolerance (Thompson, 2003, p.1), and for workplace and business success, to mention just a few life domains. The claims that IL is important, particularly in educational contexts, incites a need to measure it, both for educational purposes (e.g. assessments and documentation of student learning), and for research purposes (e.g. to evaluate the effects of educational interventions designed to further it, or to document the specifics of how IL produces the claimed benefits).

At the same time, the concept of IL is itself hard to pin down and seems to be in a state of constant flux. IL scholars and societies have published a number of complex, and at least on the surface, rather different definitions, standards, and frameworks over the last decades (e.g. Association of College and Research Libraries [ACRL], 2000, 2015; Bruce et al., 2006; Bundy, 2004; Coonan & Secker, 2011; SCONUL Working Group on IL, 2011). This introduces challenges for the measurement of IL. If IL is conceived of as a coherent, unitary, and relatively

universal construct, with a stable set of "core" features, then it should be possible to develop IL test scales (possibly with subscales) with strong internal consistency, and a structure that reflects the construct's core features. If, on the other hand, IL is a collection of rather disparate and not necessarily correlated components, as asserted by Beile O'Neil (2005, p.51), then this would have two important implications for test development. Firstly, IL tests should then be considered indexes rather than scales, as the latter measure only unidimensional constructs (Streiner, 2003). Secondly, the use of internal consistency measures, such as Cronbach's alpha, would be inappropriate measures of reliability since they rest on an assumption of unidimensionality.

Another question that is seldom addressed in the IL literature is the degree to which what we *know* about IL is reflected in what we *do* when we evaluate and use sources. Although many studies address *either* IL-knowledge *or* IL-practice, few relate them to each other.

The research described in this report therefore serves a three-fold purpose: First, to develop IL-measures for higher education that are applicable across academic disciplines, and that are brief and easy to administer, but still likely to be reliable and to support valid interpretations. Second, to determine whether what students *know* about IL corresponds to what they actually *do* when evaluating and using sources. Thirdly, to help illuminate the question of whether IL, or its source-related facets, should be conceived of as coherent, unitary constructs, or sets of disparate and more loosely related components.

## 1.1 Definitions and frameworks

IL has had many different definitions since its inception. While specific skills were highlighted in the 1980's (see American Library Association, 1989, par.3), more recent definitions are more abstract and less skill-oriented (see Association of College and Research Libraries, 2015, p.3). The working definition at the foundation of this study combines the concrete and the abstract of previous definitions: 'Information literacy encompasses the knowledge, skills and attitudes needed to be able to discover, evaluate and use information sources effectively and appropriately in order to answer questions, solve problems, create knowledge and learn.' This definition is suitable not only in educational sectors, but also in broader contexts, such as in the workplace, in daily life, and as a responsible citizen. The original, core dimensions of IL, which are as relevant today as in the term's infancy, are easily recognizable in this new definition. Our definition, and this study, emphasises the source-related core aspects of IL – finding, evaluating and using information.

Evidence of the continual evolution of the IL construct is also reflected in various IL standards and frameworks, such as those from the ACRL. Changes between the ACRL's (2000) *IL Competency Standards for Higher Education* [Standards] and its 2015 *Framework for IL for Higher Education* [Framework] reflect rapid developments in both the HE-sector and the information ecosystems in which we otherwise live and work. The amount, complexity and form of information has increased in the last 15 years, making it more difficult to navigate, evaluate and use ethically. In addition, teachers and librarians have greater responsibilities in teaching core concepts of IL and integrating them into curricula, and students have become information creators to a greater extent than previously (ACRL, 2015). Executive director Mary Ellen Davis further explains:

One of the primary reasons [for shifting to the ACRL Framework] was that the Standard was seen to reduce learning to a checklist which did not fit well with higher learning concepts that faculty were teaching. The Framework provides a bridge to the faculty for more integrated teaching of the concepts. (personal communication, Feb. 20, 2019)

In addition to the ACRL Standards (2000), other IL standards/frameworks/curricula from before 2015 (including Bundy, 2004; Coonan & Secker, 2011; SCONUL Working Group on IL, 2011),

also rely on a checklist of learning outcomes and/or skills. Although less popular after 2015, especially in the US, one advantage of the checklist approach was that it facilitated the assessment of IL competencies, in comparison to the ACRL Framework from 2015, whose abstract «threshold concepts» are harder to measure than concrete learning outcomes (Gross, et al., 2018). A notable feature of all these guides is the broad range of knowledge and skills that are incorporated into the IL concepts.

## 1.2 Existing IL knowledge tests

There are a number of different ways to assess students' IL. Library instruction and IL teaching is commonly evaluated by asking students to self-report on their learning experiences and/or subjectively estimate their IL (Oakleaf, 2008; Schilling & Applegate, 2012). Such measures, while providing important information, do not adequately capture IL knowledge, skills and behaviour, in part because students tend to overestimate their IL (Gross & Latham, 2012; Ivanitskaya et al., 2006; Julien & Hoffman, 2008; Nierenberg & Fjeldbu, 2015; Oakleaf et al., 2011; Polkinghorne & Wilton, 2010).

Measures of knowledge, usually multiple-choice tests, are another common way to measure IL (see Hollis, 2018; Mahmood, 2017; Walsh, 2009 for reviews). These are less reliant on students' more or less accurate self-assessments, and they tend to be relatively easy to administer and score. Despite their popularity, relatively few of these measures have been psychometrically evaluated (Mahmood, 2017), and some of the more thoroughly evaluated measures are only commercially available and, therefore, less accessible to most IL practitioners.

Evaluations of freely available, general IL knowledge tests typically present validity evidence in the form of documentation of test construction processes that are intended to ensure correspondence between test items and the IL standards or frameworks upon which test development was based (see Beile, 2005; Hollis et al., 2019; Podgornik, et al., 2015 for examples of this approach). To the extent that such frameworks represent consensus on the content and structure of the IL concept, this type of evidence is rightly emphasised.

In designing psychometric tests, factor analyses are often used to explore or identify any latent, underlying variables, called *factors*, from large numbers of observed variables. These factors can then be used to reduce the number of items to those reflecting the major constructs being measured. Interestingly, though, factor analytic investigations of response patterns to IL knowledge tests are rare. We have been able to locate only five. Among these, two find no clear factor structure (Beile O'Neil, 2005; Morley, 2014) and most are done on tests developed to either capture only one circumscribed sub-domain of IL (Catalano, 2015 - source evaluation; Ondrusek et al., 2005 - accessing information in a library context), or to target specific student populations (Beile O'Neil, 2005 – teacher students; Leichner et al., 2014 – psychology students; Morley, 2014 – resident physicians). Furthermore, many of these studies have methodological flaws (e.g. low item to sample ratio, no reports of assumptions checked or rationales for extraction and rotation methods) that weaken them as bases for inference. Lacking any empirical evidence of IL dimensionality is a notable gap. Of course, even though empirical analyses should not be allowed to *determine* our construct definitions, factor analyses can inform the decisions we make regarding how a given construct can or should be conceived (Mulaik, 2009; B. Thompson & Daniel, 1996). Specifically, if one assumes that IL (or the various facets of IL), is a latent variable that causes certain measurable behaviours (see Cronbach & Meehl, 1955), then factor analyses can be used to design a test structured around highly correlated items. If, on the other hand, IL is assumed to be a construct that merely represents a collection of various, possibly quite unrelated indicator behaviours, then factor analyses of test items that represent the concept should not be expected to yield a clear and stable underlying structure.

Another peculiarity of the literature on the development and evaluation of IL tests is the reliance on internal consistency measures of reliability, in particular Cronbach's alpha. Of the 16 studies reviewed by Mahmood (2017), 14 report alphas, while only four report temporal consistency (i.e., test-retest reliability). Given the aforementioned lack of factor analytic evidence regarding the dimensionality of IL tests, and the seemingly pervasive construal of IL as a multifaceted construct in extant definitions and frameworks, this is potentially problematic. If factor analytic evidence, currently lacking for most IL tests, should support unidimensionality (or, alternatively, a clear subscale structure), internal consistency measures make sense. If not, they are inappropriate. In fact, if an IL test is intended to measure a construct that consists of a set of rather different aspects, a high alpha may indicate that the test is too narrowly focused (Streiner, 2003, p.220). Such measures are best considered indexes (whose items measure behaviours that reliably cause the score), rather than scales (which assume an underlying property that reliably causes the behaviours measured by the items and that accounts for their interrelatedness).

Since the assumption of internal consistency may not apply to tests of IL, a more appropriate estimate of reliability may be their stability over time. Such temporal consistency is measured by comparing test results from the same person at different times (test-retest reliability). A reliable instrument should produce similar results if a) there is no intervention between tests, and b) the time interval is suitable. According to Streiner and Kottner (2014, p.1974), 'there should be sufficient time between the two [tests] so that the respondents do not remember their original responses and simply recall those, but not so much time that the construct being assessed can change.' Test-retest reliability is often measured with intraclass correlation coefficient (ICC, Streiner & Kottner, 2014), especially when unidimensionality is not assumed (Thorsen & Bjorner, 2009, p.31). ICC levels are considered adequate if the lower bound of the 95% confidence interval is greater than 0.6 (Multon, 2012).

### 1.3 The relationship between knowing and doing

Although several IL studies compare students' self-assessments with their competencies (Gross & Latham, 2007, 2013; Smith et al., 2013), there is little IL research attempting to establish a direct connection between their objectively measured IL-knowledge (knowing) and applied skills (doing) in an authentic educational context. Two exceptions are work by Schilling and Applegate (2007) and Beile O'Neil (2005), which both focus on information seeking. Neither of these, however, measure applied skills in actual, mandatory assignments. Although Schilling and Applegate (2007) found that students' IL-knowledge (as measured by an objective test) was *not* related to their subsequent performance on a practical literature search task, Beile O'Neil (2005) found in a similar study that there was a significant correlation between the two, although her sample size was small.

### 1.4 The present study

Three quantitative measures of IL were employed in order to answer the research questions in the present study, and this article describes both the development and testing of these measures, and the implications of the results. The first is a *know*-measure, in the form of a test, and the other two are *do*-measures, based on mandatory student assignments:

1. Know: a multiple-choice IL knowledge test to capture knowledge of three key, source-focused aspects of IL (seeking information, evaluating and using sources)
2. Do: a source evaluation measure to capture students' abilities to select and critically evaluate sources
3. Do: a source use measure to capture students' abilities to use sources correctly when writing



While the multiple-choice knowledge test (hereafter “IL knowledge test”) measures what participants *know* about the source-focused components of IL, an annotated bibliography and rubric measure what they *do* in practice when evaluating and using sources, respectively.

The present study differs from previous know-do research in two important ways. Firstly, while Schilling and Applegate (2007) and Beile O’Neil (2005) focused on information seeking, the present study examines associations between knowing and doing in terms of the evaluation and use of sources. Secondly, the present study uses authentic results from mandatory coursework instead of non-graded tasks designed purely for research purposes. Establishing a connection between what the students *know* in theory and what they *do* in practice can be used to test the validity of the IL knowledge test as an indicator of practical abilities related to evaluating and using sources.

Since these tools were designed to measure IL knowledge and skills in higher education (HE), students at different levels were recruited to test their efficacy. An additional survey, measuring respondents’ interest in being/becoming information literate individuals (what they *feel*), will be addressed in another article.

In order to explore the possible structure of the IL knowledge test, exploratory factor analyses were undertaken in this study. Given that items were developed to capture three core, source-focused aspects of IL, one might expect extracted factors to align with these aspects. On the other hand, if IL is conceived of as ‘one underlying variable’ or ‘singular construct’ (Hollis, 2018, pp.76–77), a latent attribute that causes the observable behaviours that IL tests measure, then one would expect a strong first factor, with remaining factors accounting for very little variance (Streiner, 2003). A third possibility is that a broadly conceived IL test measures a number of rather different facets that may or may not be related to each other in any other way than that they tend to be seen as part of a family of related but distinct IL skills, in which case we should not expect any clear or meaningful factor structure. Given that all of these are reasonable a priori possibilities, exploratory factor analysis seemed the most appropriate analytic approach.

The present study aims to create tools, suitable for multiple disciplines in HE, for measuring IL’s source-components, and to psychometrically evaluate these with large sample sizes. The study will also compare what students *know* with what they *do* when evaluating and using sources for real course work. It will also examine the dimensionality of IL, with possible implications for how it should be measured.

## 2. Methods

### 2.1 Participants

Data was collected from four different samples, serving slightly different but overlapping purposes. The study was approved by the Norwegian Data Protection Authority, and informed consent was required for participation.

First, for the item selection and wording refinement process, four selected IL experts were recruited to evaluate the validity of the 50 original items. These included a senior lecturer at The Swedish School of Library and Information Science (LIS), and three IL-practitioners in Norway, including the head of the library’s Department of Public Services at UiT The Arctic University of Norway, and an assistant professor of education and LIS, and an academic librarian at Oslo Metropolitan University.

Five students participated in think-aloud protocols while answering the test questions. Three were first-year bachelor students and two were college-bound high school seniors (four female and one male). The last two are (or will soon be) new college students – the least experienced

of the survey's main target group (undergraduates) – and therefore more likely than higher-level students to have questions/misunderstandings regarding items.

A pilot sample (Table 1) of undergraduates were then recruited in spring 2019 through their Learning Management Systems (LMS). Of these respondents, 40% were in the first year of their current study program. Two universities, with academic studies such as psychology and economics, and one university of applied sciences (UAS), with vocational studies such as nursing and teacher education, were included. All are medium-sized (14,000-18,000 students) public institutions with free tuition. This sample answered all of the 50 initially created items for the IL knowledge test. The data were intended primarily for item selection and test construction purposes.

**Table 1:** Descriptive statistics: pilot sample – undergraduate (survey in Norwegian)

<i>N</i>	Semesters completed	
	<i>M</i>	<i>SD</i>
268	3.5	2.58

The first students to complete the final version of the 21-item IL knowledge test were undergraduates at the beginning of their first semester, *before* IL instruction, in 2019 (Table 2). These students had previously completed an average of 0.8 semesters of HE in other disciplines. Respondents were from three universities and one UAS, from south-eastern to northern Norway, in academic and vocational fields ranging from sociology to fisheries management. The institutions are public and small- to medium-sized (5,200-18,000 students). Participants were recruited either in the classroom, via their LMS, or by teaching-librarians from their university libraries, and had the possibility of winning a gift card. The data were intended for assessment of reliability and validity, and for exploring any potential factor structure in the IL test. Many participants (37%) were in their first year of a psychology program at the largest university in our sample. These psychology students provided data for both the IL knowledge test and the two do-measures, by utilizing assignments in their academic skills course.

The undergraduate sample answered the IL knowledge test once again at the end of the semester, *after* IL instruction. They were recruited via e-mail and offered a possible gift card for their participation. During the semester, all received instruction on information searching, and the evaluation and use of sources.

**Table 2:** Descriptive statistics: undergraduate sample – beginning and end of first semester (survey in Norwegian)

Group	<i>n</i>	Gender			Age ( <i>M</i> )	Semesters in HE	
		Male	Female	Other		<i>M</i>	<i>SD</i>
<b>Beginning of first semester</b>							
Psychology students	97	31	66	0	21.3	0.5	0.89
Other students	163	65	98	2	23.4	1.0	1.15
Total	260	95	163	2	22.8	0.8	1.09
<b>End of first semester</b>							
Psychology students	45	14	31	0	20.6	0.6	1.47
Other students	67	16	50	1	23.1	1.5	2.06
Total	112	30	81	1	22.1	1.1	1.89

A graduate student sample of master's and PhD students in a variety of disciplines, from several Norwegian and five international universities, were recruited via social media and email from faculty at other institutions (Table 3). This sample answered the IL knowledge test only once and provided a comparison group for the undergraduate sample.

**Table 3:** Descriptive statistics: graduate student sample

	<i>n</i>	Gender			Age ( <i>M</i> )	Language	
		Male	Female	Other		Norwegian	English
Master's	196	57	131	5	32.3	80	116
PhD	170	69	96	2	33.4	65	105
<b>Total</b>	366	126	227	7	32.7	145	221

## 2.2 Materials

### 2.2.1 IL knowledge test

#### 2.2.1.1 Item generation

Before generating items for the IL knowledge test, ten international IL frameworks and standards were analysed in order to find one that: a) applies to most disciplines in HE, and b) specifies learning outcomes relevant to central constructs of IL identified in the framework analysis, thereby facilitating assessment. The framework that best fulfilled these criteria was the Australian and New Zealand IL Framework [ANZIL framework] (Bundy, 2004), based on the ACRL's now discontinued Standards (2000). Despite its age, the decision was therefore made to base IL knowledge test items on the ANZIL framework's six standards, regarding respectively: recognizing the need for information, and finding, evaluating, managing, creating, and using information (see Bundy, 2004).

A pool of 50 items (Nierenberg et al., 2021, Pilot) was created with the goal of measuring students' knowledge of aspects of IL described in the framework. Items had varying levels of difficulty and were intended to detect learning over time. All items were multiple choice with four alternative, plausible answers, one of which was correct. 'I do not know' was not an option. An example of a survey item is: 'What is the most important reason to use sources when writing a paper? a) To support arguments; b) To avoid plagiarism; c) To show that you've read the sources; d) To satisfy the requirements of the assignment.'

As discussed below, items were selected from the original pool of 50 items based on expert evaluations of validity and responses to the items in a pilot sample.

#### 2.2.1.2 Expert evaluations of validity

The panel of experts looked specifically at three qualities (headings from Beile, 2005):

1. Clarity: Are items and answer alternatives clear, unambiguous, and understandable?
2. Content accuracy: How well does each item align with the standards and learning outcomes in the ANZIL framework?
3. Objectivity and bias: Are items objective? Are there cultural, geographic, disciplinary, or other biases? (The survey is designed to be used in various contexts and suitable for *all* students, independent of their field of study, location, race, gender, religion, sexual orientation, etc.)

Items were rated by the experts on a scale of 0 (absence of quality) to 3 (fully expressed quality), and those items receiving a score of 2.6 or below were later revised in accordance with experts' feedback.

### 2.2.1.3 Pilot test

Responses from the pilot sample ( $n = 268$ ) to the 50 original items were analysed to select those most useful for the final survey. In psychometrically designed tests, items should: a) have a suitable range of difficulty, and b) correlate positively to the total test score, i.e., item discrimination (DeMars, 2017). Firstly, items were discarded if more than 88% of the sample answered correctly. Secondly, items with an item-to-total correlation (the point-biserial correlation between the item score and the summed test score) lower than 0.3 were removed. After deleting 38 items based on these criteria, 22 items remained.

Interestingly, these 22 items were all related to information seeking and to the evaluation and use of information sources, thus corresponding to Standards Two, Three and Six in the ANZIL framework (see Bundy, 2004). Items from the other three standards (recognizing the need for information, managing information, and creating new information) were not useful based on the criteria for item difficulty and discrimination. The 22 remaining items were then grouped in terms of difficulty, measured by the percentage of participants who answered the item correctly (P-value): a) easy (P-value 80% - 88%), b) medium (P-value 50% - 79%), and c) hard (P-value < 50%). In cases where multiple items at the same level of difficulty covered the same IL topic, for example several easy items about source evaluation, items with the lowest item discrimination were deleted. When this process was complete there were seventeen items remaining. These items thus demonstrate two characteristics: a) a range of difficulty (without being too easy, to avoid ceiling effects), and b) at least a moderate correlation with total test scores. These seventeen items should therefore be useful in distinguishing between respondents with varying levels of IL.

Four additional items were added to the test at this stage in order to have an equal number of items in the three main categories of information seeking, and the evaluation and use of sources. The final IL knowledge test now consisted of 21 items, four of which had neither been tested in the pilot nor rated by experts.

Since we wanted to make a tool that could be used in multiple languages, the final IL knowledge test was then translated into English, and we tested both language versions with the master's and PhD sample. The original Norwegian version and the English translation are provided in Nierenberg et al. (2021).

Exploratory factor analyses based on data from the pilot sample were also performed at this stage, with the intention to further refine the item selection (see Appendix A) and explore dimensionality of the IL test.

### 2.2.1.4 Think aloud protocol

In order to test the remaining survey items for readability and comprehension, think-aloud protocols were collected from five current or prospective students. In a think-aloud protocol, members of the target population verbalise their thought processes while performing an action, in this case while taking a survey, saying aloud how they interpret the question and choose their answer (Hughes, 2018, p.767). This can uncover possible misunderstandings, confusion, or ambiguities that were not previously detected. Although think-aloud protocols produce useful evidence for establishing psychometric accuracy, they are nevertheless seldom utilized (Hughes, 2018, pp.764–765).

While taking the IL knowledge test, participants were encouraged to describe verbally how they interpreted the questions, chose their answers, and eliminated answer alternatives. They could also comment on the test in general afterwards. Participants used 15 to 25 minutes to take the survey while thinking aloud. The wording in four items was revised as a result of their feedback.

### 2.2.2 Annotated bibliography for source evaluation

One of the tools employed to measure what students actually *do*, (as opposed to the IL knowledge test which measures what they *know*), is the annotated bibliography for source evaluation. Several items in the IL knowledge test address the critical evaluation of sources, but in order to see if students actually *use* criteria for source evaluation in practice, an authentic, assignment-based method was desired.

When reviewing the literature, little research was found on measuring source evaluation with quantitative methods; one exception was work by Walton and Hepworth (2012). Their method allows qualitative data from students' annotated bibliographies, in which they evaluate their sources of information, to be assigned quantitative values (Walton & Hepworth, 2012, p.58). A variation of this method was used in the present study. For one of their mandatory assignments, the psychology students from the undergraduate sample chose three academic sources for their term papers and composed references to these sources in a bibliography. Their choice of sources was based on what they had learned about source evaluation during the semester. For each reference, students wrote an annotation describing why it was a good, academic source, and well suited to answering their research question. For the purposes of the present study, annotations from those students who had previously completed the IL knowledge test at the start of the semester ( $n = 93$ ) were analyzed and rated for the following three components:

1. *Quality* of the source: ratings of each source on a scale of 0 (not academic) to 3 (academic).<sup>1</sup>
2. *Variety* of criteria: the number of different, unique criteria for source evaluation stated by the student (e.g. relevancy, accuracy, authority, etc.) in each annotation. Each specific criterion is allocated a score of '1'.
3. *Frequency* of criteria: the number of instances criteria for source evaluation appear in each annotation. Each criterion is allocated a score of '1'. (This will be different from the *variety* score if a criterion, e.g. authority, is listed more than once in the annotation.)

Details about the annotated bibliography measure for source evaluation, with an example annotation, are provided in Nierenberg et al. (2021). All bibliographies were scored by two raters – one of this article's authors and a master's student in psychology. To assure the reliability of their scoring, the raters compared, discussed, and agreed upon assessments for the first ten students' annotations and independently scored the rest. During this process, raters created a list of eligible, "countable" criteria for source evaluation in order to make scoring as objective as possible. Inter-rater reliability for the raters was assessed using ICC when scoring was complete. ICC is the best measure of rater consistency for interval data with a normal distribution, and values over 0.60 are considered acceptable (Bujang & Baharum, 2017; Multon, 2012).

### 2.2.3 Rubric for source use

After completing the IL knowledge test and the source evaluation assignment, the remaining psychology students ( $n = 87$ ) wrote a term paper on a psychology topic of their choice. In this mandatory assignment, students were required to use a minimum of five academic sources and to cite their sources in APA-style. The assignment was graded using a rubric with 17 evaluation

---

<sup>1</sup> This is based on the following definition of an academic source used in the introductory psychology course: a) An academic source is written by professionals, for professionals. b) It builds on other academic sources, cites sources to substantiate claims and arguments, and provides complete references for these sources. c) The text has been the subject of a thorough, professional editorial process and has received constructive criticism before being published.

criteria, five of which pertained to the use of sources (see Table 4, and Nierenberg et al., 2021) for more details).

**Table 4:** Criteria and scoring in rubric for assessing students' use of sources

Criteria for use of sources	No	Partially	Yes
Are academic sources used to support arguments?	0	0.5	1
Are sources cited in the text when necessary?	0	0.5	1
Are the in-text citations written in correct APA-style?	0	0.5	1
Is the reference list written in correct APA-style?	0	0.5	1
Are all in-text citations listed in the reference list, and vice versa?	0	-	1

These five criteria are designed to evaluate the students' abilities to properly use and cite information sources when writing. Three raters, two of this article's authors and a master's student in psychology, evaluated the term papers of the 87 students who had previously answered the IL knowledge test, using this 5-point source use rubric. Raters calibrated their grading beforehand to ensure reliability, and then independently scored the rest. Though APA was our reference style of choice, any reference style can be substituted in its place.

## 2.3 Reliability and validity evidence

Analyses were conducted in IBM SPSS Statistics 26 (2019) and RStudio (2019).

### 2.3.1 Reliability analyses

Temporal consistency, measured with ICC test-retest reliability, was calculated from data from 46 undergraduates who took the IL knowledge test twice at the end of the semester. Time intervals varied between 7 and 73 days, with no IL instruction in between, an acceptable test-retest range (Streiner & Kottner, 2014). Inter-rater reliability was assessed for consistency using ICC for both of the do-measures.

### 2.3.2 Validity analyses

In addition to the comprehensive item selection and refinement process based on pilot testing, expert review and student think-aloud protocols, further support for the validity of the final IL knowledge test would be based on results showing that the overall knowledge test score discriminates among students at different levels of HE, and between undergraduates at the start and end of one of their first semesters. To detect these differences, two analyses were performed. Since the assumption of homogeneity of variance was not met for this data, we used a Welch's ANOVA with Games-Howell post hoc to compare mean IL knowledge test scores for undergraduates (semester-start), master's and PhD-students. In addition, a paired samples t-test on test scores at the start and end of the semester for undergraduate students was performed.

Bivariate correlations were then calculated to find potential relationships between scores on the IL knowledge test and scores on do-measures. Such correlations would provide evidence that what students *know* is related to what they *do* in practice.

### 2.3.3 Exploratory factor analyses

Given that we constructed items to reflect what might be considered core dimensions of IL, we wanted to explore a possible structure of the IL knowledge measure. We did this not only to aid in item selection, but also to explore whether hypothesised or postulated dimensions of IL were

revealed in the structure of the correlations among items. A number of factor models were explored, based on matrices of tetrachoric correlations between items. Factors were extracted using principal axis factoring, and solutions were rotated using ProMax. Details of these analyses are described in Appendix A.

### 3. Results

All data used in this article is available from UiT Open Research Data (Nierenberg et al., 2021).

#### 3.1 Know: IL knowledge test

Mean scores, standard deviations, and minimum and maximum scores for the IL knowledge test for the different levels of HE are shown in Table 5. For score distribution histograms, see Appendix B.

**Table 5:** IL knowledge test score: statistics for student groups

Student group	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
Undergrad. (semester start)	260	12.46	2.95	3	19
Undergrad. (semester end)	112	13.92	2.97	6	20
Master's	196	16.06	2.49	8	21
PhD	170	16.82	2.11	10	21

##### 3.1.1 Exploratory factor analyses

Factor analyses with pilot data provided no useful basis for decisions regarding item selection. Correlation matrices showed generally poor factorability in all samples, and none of the many potential solutions were readily interpretable. Specifically, neither a one-factor solution, which one might expect if items of the IL knowledge test represented a unitary underlying construct, nor solutions where factors aligned with core facets of IL, or indeed any other meaningful pattern, emerged from these analyses (see Appendix A)

##### 3.1.2 Reliability evidence

Reliability statistics for the IL measures, based on the undergraduate sample, are shown in Table 6. Intraclass correlation coefficients for consistency – for both interrater reliability and test-retest reliability – are all above .80, normally considered sufficient evidence of reliability. Results from graduate students show no difference in mean scores between the Norwegian ( $M = 16.59$ ,  $n = 145$ ,  $SD = 2.33$ ) and English ( $M = 16.29$ ,  $n = 221$ ,  $SD = 2.35$ ) versions of the IL knowledge test ( $t_{364} = -1.212$ ,  $p = .616$ ). These data are therefore pooled in all subsequent analyses.

**Table 6:** Reliability statistics for IL measures

IL measure	Reliability measure	Unit	Value	<i>N</i>	95% CI	
					LL	UL
IL knowledge test (21 items)	test-retest reliability	ICC	.84	46	.72	.91
Source evaluation measure	interrater reliability <sup>a</sup>	ICC	.89	93	.83	.93
Source use measure	interrater reliability <sup>b</sup>	ICC	.92	87	.89	.95

*Note.* ICC = intraclass correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

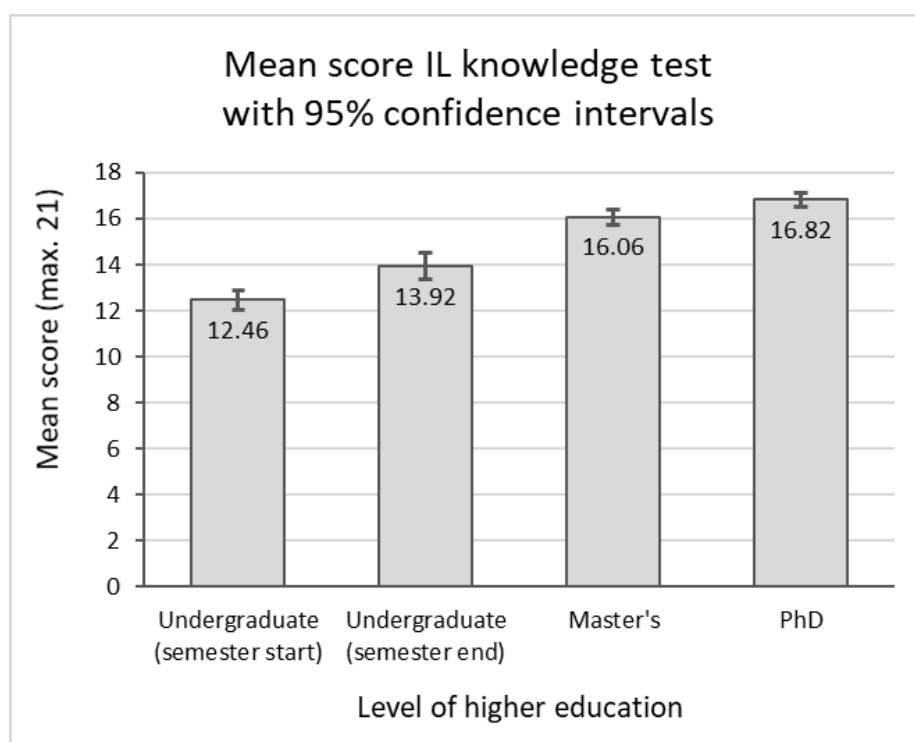
<sup>a</sup> 2 raters

<sup>b</sup> 3 raters

### 3.1.3 Validity evidence

The previously described processes of item generation and item selection provide validity evidence for the final IL knowledge test items. In addition, a paired sample t-test for undergraduates' test scores at the beginning and end of their first semester showed a statistically significant difference ( $t_{111} = 5.078, p < .001$ ). On average, semester-end scores were 1.21 points higher than semester-start scores (95% CI [0.74, 1.69]) for those students who took both tests.

A Welch's ANOVA showed statistically significant differences between the means of undergraduates (semester-start), master's and PhD scores,  $F(2, 623) = 179.49, p < .001$ . Follow-up procedures, using the Games-Howell post hoc, indicated that the mean score for the undergraduate group ( $M = 12.46, SD = 2.95$ ) was significantly different ( $p < .001$ ) from the master's group ( $M = 16.06, SD = 2.49$ ), and the mean for the master's was significantly different ( $p = .005$ ) from the PhD group ( $M = 16.82, SD = 2.11$ ). Taken together, these results suggest that the IL knowledge test discriminates between students at different levels of HE – more advanced students have higher mean scores. Mean scores with 95% confidence intervals are plotted in Figure 1.



**Figure 1:** Mean score IL knowledge test with 95% confidence intervals for different levels of HE

### 3.2 Reliability and validity of assignment-based measures

The interrater reliability for the scoring of both *do*-measures was high (see Table 6). Both measures have inherent validity since teachers devised them to measure aspects of students' IL.

### 3.3 Know and do: relationship between the IL-test and do-measures

IL knowledge semester-start test scores from psychology students in the undergraduate sample were used in correlation and regression analyses. The correlation matrix and details of regression analyses are found in Appendix C.



### 3.3.1 Source evaluation measure: annotated bibliography

No significant correlation was found between annotated bibliography total scores and IL knowledge test scores. However, when correlating annotated bibliography component scores for quality, variety and frequency with test scores, a statistically significant, positive, but weak/moderate correlation was found between the quality component score and test score,  $r(93) = .27$ ,  $p = .008$ . Test scores were not significantly correlated with variety or frequency component scores.

In an attempt to control for HE experience, multiple regressions were performed, using test scores and number of semesters of higher education as predictors, and the quality component score as dependent variable. Unsurprisingly, given the restricted range of the number of semesters variable and its lack of association to the quality score as measured by the simple correlation, it does not influence the predictive strength of the knowledge test scores ( $\beta = .27$ ,  $t(91) = 2.70$ ,  $p = .008$ ). (See Appendix C for details.)

### 3.3.2 Source use measure: rubric

Bivariate correlation analyses with source use rubric scores indicate a statistically significant, positive, and moderate/weak correlation with IL knowledge test scores,  $r(87) = .31$ ,  $p = .004$ . Again, we did multiple regressions, using test scores and number of semesters of higher education as predictors, this time using source use rubric score as the dependent variable. This yields a  $\beta = .32$ ,  $t(85) = 3.06$  for the test scores when number of semesters is accounted for (see Appendix C).

## 4. Discussion

The goals of this study were to 1) develop quantitative IL assessment measures for HE that support valid interpretations and have a high degree of reliability, 2) provide these in two languages (Norwegian and English), 3) determine whether there is a connection between what students *know* about IL and what they *do* in practice when evaluating and using sources, and 4) shed light on the question of whether the source-related components of IL can be considered a construct with clear dimensional structure.

To answer these research questions, three measures were developed and examined: a test to measure IL knowledge and two assignment-based tools to measure what students do in practice with their IL knowledge when evaluating and using sources in mandatory coursework. In contrast to self-report IL-measures, which are commonly used to evaluate the impact of IL instruction, the instruments in the present study provide more objective measures of students' IL knowledge and skills.

### 4.1 Assessment measures

#### 4.1.1 Know: IL knowledge test

The IL knowledge test presented in the current study can provide a valuable contribution to IL-assessment and has several advantages over many other IL tests (see Mahmood, 2017), as demonstrated by the following:

- A thorough item selection process based on framework analysis, pilot testing, expert evaluations and think-aloud protocols, ensures that the IL knowledge test covers three core aspects of IL that are common to nearly all of IL's many definitions/interpretations.
- It has been evaluated for reliability and validity in a number of different samples.
- It is freely available, and quickly and easily administered.
- It is suitable for students in diverse disciplines.

- Reliability was estimated using temporal consistency (test-retest reliability), as opposed to by calculating internal consistency (as most IL test developers do, despite its relevancy only for unidimensional constructs).
- Scores on the IL knowledge test are nearly identical in the English and Norwegian versions, suggesting that it can be used in both languages.
- It discriminates reasonably well between undergraduate students at the beginning and end of their first semester, and between students at different levels of HE.

That only items regarding the source-related components of IL were found useful in measuring IL knowledge was an unexpected result of the item selection process. This conformed with our intent to base the IL knowledge test on those aspects of IL common to its various definitions, and underscores that finding, evaluating and using information are fundamental to the concept of IL.

#### **4.1.2 Do: assignment-based measures**

The source evaluation and source use measures utilized in the present study can prove valuable both in quantitatively assessing two core aspects of IL – the evaluation and use of sources – and in examining the connection between knowing and doing in an educational context. Both measures can be scored with high interrater reliabilities. Arguably, they have inherent validity, since they were devised by faculty to assess aspects of students' IL explicitly valued in a real HE context. Importantly, they directly assess performance on practical IL tasks, rather than «mere» knowledge. While there is some evidence that knowing and doing are related (see section 4.2), the two are clearly not equivalent.

Results indicate that only the quality component of the source evaluation measure (ultimately the most important) has significant correlations with IL knowledge test scores at the start of students' first semester. The relevance of source evaluation variety and frequency is worthy of additional scrutiny with other student populations or at different points of the HE experience. Still unknown is if student low variety and frequency scores are a result of an inability to discern how they arrive at their quality appraisal, or their inability to articulate this.

## **4.2 Knowing and doing**

Our study complements the paucity of IL literature regarding the connection between IL knowledge and skills in areas other than information seeking. The present study differs from similar research, for example Beile O'Neil (2005) and Schilling and Applegate (2007), by including two core aspects of IL – the evaluation and use of sources – and by assessing students' skills using authentic, graded coursework. Results provide varied evidence of the connection between what students know and actually do. Scores on the IL knowledge test had positive, significant, and borderline weak-moderate correlations with source evaluation and source use scores. This may be due to the students' short experience with HE and indicate that those who are early in their academic careers are not yet able to effectively use their knowledge in practice. A future study will measure these correlations toward the end of the students' undergraduate program to determine how the ability to translate knowledge into practice evolves with HE experience.

The modest strength of these associations indicates that students' demonstrated practice is strongly influenced by factors other than their knowledge of IL. Assessments of students' IL should therefore be based on multiple measures covering both knowledge and practical performance, and not exclusively on IL knowledge tests. We attempted to control these associations for amount of HE experience, however the range of this variable was rather restricted in our sample, so the control was weak. While amount of HE nevertheless correlated with IL test scores, it did not correlate with any of the do-measures in our first-semester student sample.

### 4.3 Dimensionality

There is little consensus among IL researchers about the dimensionality of the IL construct, although its homogeneity is often implied. Few studies, with the exception of Beile O'Neil (2005) and Morley (2014), have investigated the dimensionality of IL as a whole using factor analyses, and both concluded that IL is multidimensional. Hollis (2018), on the other hand, explicitly postulates that IL is a unidimensional construct, but has yet to provide factor analytic evidence of this.

Rather than investigate IL as a whole in this study, we have focused on its source-related components – finding, evaluating and using information. Factor analyses of the IL knowledge test show no interpretable factor solutions, and thereby provide little evidence of the unidimensionality of these components, either as a group or individually (with one or several identifiable factors respectively). Contrary to some previous factor analytic investigations of freely available IL tests (Leichner et al., 2013; Morley, 2014), this cannot be dismissed on the grounds of insufficient sample sizes. Since the assumption of unidimensionality is not fulfilled, it is not appropriate to use either Cronbach's alpha or other measures of internal consistency for the IL knowledge test.

These findings thereby parallel Beile O'Neil's (2005, p.51) conclusions in her study of the broader IL construct. With no clear factor structure, test items should not be expected to exhibit cohesion. Students can possess knowledge of some concepts of information seeking without having an understanding of other concepts. There is nothing inherent about, for instance, knowing how to use the Boolean 'OR' and knowing how to choose appropriate databases or keywords for your search. Consequently, IL tests, to the extent that they measure heterogeneous constructs, should be considered indexes, rather than scales (cf. DeVelis, 2017; Streiner, 2003), and their reliability should not be assessed using internal consistency measures. One advantage of considering the test an index rather than a scale is that items measure different aspects of IL knowledge and can thus provide valuable formative assessment information over and above the total score on the test.

### 4.4 Limitations

The four items (18-21) not tested in the pilot may not meet the requirements for item difficulty and discrimination used in item selection. While post hoc testing indicates that all four are within the appropriate range of difficulty for undergraduates, items 18 and 19 may not contribute sufficiently in distinguishing between students with different levels of IL ( $r < .30$ ).

Although reliability and validity for the IL knowledge test and assignment-based measures have been assessed in this study, they should not be considered fully validated for general use, as reliability and validity evidence ideally are established for each specific sample and circumstance (American Educational Research Association, 2014; Streiner & Kottner, 2014).

### 4.5 Recommended uses of the IL measures

The three measures in this study are each suitable for the aggregated assessment of groups of students for research purposes, and to assess changes in IL levels as a result of interventions. In the latter case, this assumes that the content of the interventions is relevant to information seeking, and source evaluation and use. They may also serve as support for formative assessment and as ingredients in learning activities. To achieve a more accurate evaluation of individual students, for instance for purposes of high stakes educational assessment, we recommend combining several IL measures.

The IL knowledge test is designed to measure undergraduates' knowledge of three core aspects of IL – information seeking, source evaluation and source use. Its scope is therefore

narrower than for instance the Open Test of Information Literacy (Hollis et al., 2019) and the most comprehensive commercial tests, yet broader than studies focusing on only one aspect of IL, such as Catalano (2015) and Ondrusek et al. (2005).

#### 4.6 Suggestions for future research

An idea for future research is to follow students longitudinally to assess changes in the association between knowing and doing as they progress. Although several studies have shown increases in IL knowledge *or* skills after interventions, changes in both and correlations between the two over time, measuring with authentic assignments, has not previously been documented. More varied assessment in future research, including both qualitative and additional quantitative measures, could potentially provide stronger evidence for connections between know and do than were found in the present study. We invite others to develop these measures further by following our method, and to produce new measures for other aspects of IL knowledge and practice, thereby expanding this suite of tools in valid and reliable ways.

### 5. Conclusion

The assessment measures developed and tested in the present study provide a unique combination of tools for quantitatively measuring not only knowledge of core facets of IL, but also applied skills. This does not necessarily mean that these are *complete* measures of students' knowledge and ability to find, evaluate and use sources, but that they measure at least some aspects of these core features of IL that are deemed important in a typical HE context.

We believe this study makes four important contributions. Firstly, although other measures focusing on the same core components of IL are available, few are suitable for multiple disciplines in HE and have been tested with as large a sample size. Secondly, the measures are thoroughly evaluated – also for temporal consistency – allowing reliable and valid interpretations of their results. Thirdly, the study is the first of its kind, as far as we know, to compare *knowing* and *doing* regarding the evaluation and use of sources, utilizing authentic coursework. Fourthly, based on evidence that IL is a heterogeneous construct, the study examines whether IL's source-components can be considered unidimensional constructs on their own or together, and finds little evidence of their homogeneity. We argue, therefore, that unless there is evidence of clear dimensionality, IL tests are best considered indexes rather than scales.

### References

American Educational Research Association, National Council on Measurement in Education, & American Psychological Association (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Library Association (1989). *Presidential committee on information literacy: Final report*. Association of College & Research Libraries.  
<http://www.ala.org/acrl/publications/whitepapers/presidential>

Association of College and Research Libraries. (2000). *Information literacy competency standards for higher education*. American Library Association.  
<http://www.ala.org/acrl/standards/informationliteracycompetency>

Association of College and Research Libraries. (2015). *Framework for information literacy for higher education*. American Library Association.  
<http://www.ala.org/acrl/files/standards/standards.pdf>

- Beile O'Neil, P. (2005). *Development And validation of the Beile test of information literacy for education (b-tiled)*. [Doctoral dissertation, University of Central Florida]. STARS repository. <https://stars.library.ucf.edu/etd/530/>
- Beile, P. (2005). *Development and validation of the information literacy assessment scale for education (ILAS-ED)*. [Conference paper]. American Educational Research Association Annual Meeting, Montreal, Canada. <http://eprints.rclis.org/16972/>
- Bruce, C., Edwards, S., & Lupton, M. (2006). Six frames for information literacy education: A conceptual framework for interpreting the relationships between theory and practice. *Innovation in Teaching and Learning in Information and Computer Sciences*, 5(1), 1–18. <https://doi.org/10.11120/ital.2006.05010002>
- Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Sciences*, 12(1), 1–11.
- Bundy, A. (2004). *Australian and New Zealand information literacy framework: Principles, standards and practice* (2nd ed.). Australian and New Zealand Institute for Information Literacy. <http://www.libnet.sh.cn/upload/htmleditor/File/130620025617.pdf>
- Catalano, A. (2015). The effect of a situated learning environment in a distance education information literacy course. *The Journal of Academic Librarianship*, 41(5), 653–659. <https://doi.org/10.1016/j.acalib.2015.06.008>
- CILIP Information literacy group. (2018). CILIP definition of information literacy 2018. CILIP: The Library and Information Association. <https://www.cilip.org.uk/news/421972/What-is-information-literacy.htm>
- Coonan, E., & Secker, J. (2011). *A new curriculum for information literacy (ANCIL): Curriculum and supporting documents*. University of Cambridge. <https://www.repository.cam.ac.uk/handle/1810/244638>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281. <https://doi.org/10.1037/h0040957>
- DeMars, C. E. (2017). Classical test theory and item response theory. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (Vol. 1-2, pp. 49–73). John Wiley & Sons.
- DeVelis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Gross, M., & Latham, D. (2007). Attaining information literacy: An investigation of the relationship between skill level, self-estimates of skill, and library anxiety. *Library and Information Science Research*, 29(3), 332–353. <https://doi.org/10.1016/j.lisr.2007.04.012>
- Gross, M., & Latham, D. (2012). What's skill got to do with it? Information literacy skills and self-views of ability among first-year college students. *Journal of the American Society for Information Science and Technology*, 63(3), 574–583. <https://doi.org/10.1002/asi.21681>
- Gross, M., & Latham, D. (2013). Addressing below proficient information literacy skills: Evaluating the efficacy of an evidence-based educational intervention. *Library and Information Science Research*, 35(3), 181–190. <https://doi.org/10.1016/j.lisr.2013.03.001>

- Gross, M., Latham, D., & Julien, H. (2018). What the framework means to me: Attitudes of academic librarians toward the ACRL framework for information literacy for higher education. *Library & Information Science Research*, 40(3), 262–268. <https://doi.org/10.1016/j.lisr.2018.09.008>
- Hollis, H. (2018). Information literacy as a measurable construct: A need for more freely available, validated and wide ranging instruments. *Journal of Information Literacy*, 12(2), 76–88. <https://doi.org/10.11645/12.2.2409>
- Hollis, H., Rachitskiy, M., & van der Leer, L. (2019). The development and face validity testing of the Open Test of Information Literacy with context-specific add-ons: OTIL. *LIBER Quarterly*, 29(1). <http://doi.org/10.18352/lq.10264>
- Hughes, D. J. (2018). Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp.751–779). John Wiley & Sons.
- Ivanitskaya, L., O'Boyle, I., & Casey, A. M. (2006). Health information literacy and competencies of information age students: Results from the interactive online Research Readiness Self-Assessment (RRSA). *Journal of Medical Internet Research*, 8(2).
- Julien, H., & Hoffman, C. (2008). Information literacy training in Canada's public libraries. *The Library Quarterly*, 78(1), 19–41. <https://doi.org/10.1086/523908>
- Leichner, N., Peter, J., Mayer, A. K., & Krampen, G. (2014). Assessing information literacy programmes using information search tasks. *Journal of Information Literacy*, 8(1), 3–20. <https://doi.org/10.11645/8.1.1870>
- Mahmood, K. (2017). A systematic review of evidence on psychometric properties of information literacy tests. *Library Review*, 66(6/7), 442–455. <https://doi.org/10.1108/LR-02-2017-0015>
- Morley, S. K. (2014). *Initial development of a medical information literacy questionnaire* [Doctoral dissertation, The University of New Mexico College of Education, USA]. The University of New Mexico repository. [https://digitalrepository.unm.edu/educ\\_ifce\\_etds/30/](https://digitalrepository.unm.edu/educ_ifce_etds/30/)
- Mulaik, S. A. (2009). *Foundations of factor analysis* (2nd ed.). CRC Press.
- Multon, K. D. (2012). Interrater reliability. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp.627–628). <https://dx.doi.org/10.4135/9781412961288>
- Nierenberg, E., & Fjeldbu, Ø. G. (2015). How much do first-year undergraduate students in Norway know about information literacy? *Journal of Information Literacy*, 9(1), 15–33. <https://doi.org/10.11645/9.1.1983>
- Nierenberg, E., Låg, T., & Dahl, T. I. (2021). *Replication Data for: Knowing and doing: The development of information literacy measures to assess knowledge and practice* (DataverseNO: Version V2) [Data set, survey and evaluation measures]. DataverseNO. <https://doi.org/doi:10.18710/L60VDI>
- Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy approaches. *portal: Libraries and the Academy*, 8(3), 233–253. <https://doi.org/10.1353/pla.0.0011>

- Oakleaf, M. (2014). A roadmap for assessing student learning using the new framework for information literacy for higher education. *The Journal of Academic Librarianship*, 5(40), 510–514. <https://doi.org/10.1016/j.acalib.2014.08.001>
- Oakleaf, M., Millet, M. S., & Kraus, L. (2011). All together now: Getting faculty, administrators, and staff engaged in information literacy assessment. *portal: Libraries and the Academy*, 11(3), 831–852. <https://doi.org/10.1353/pla.2011.0035>
- Ondrusek, A., Dent, V. F., Bonadie-Joseph, I., & Williams, C. (2005). A longitudinal study of the development and evaluation of an information literacy test. *Reference Services Review*, 33(4), 388–417. <https://doi.org/10.1108/00907320510631544>
- Podgornik, B. B., Dolničar, D., Šorgo, A., & Bartol, T. (2015). Development, testing, and validation of an information literacy test (ILT) for higher education. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.23586>
- Polkinghorne, S., & Wilton, S. (2010). Research is a verb: Exploring a new information literacy-embedded undergraduate research methods course. *Canadian Journal of Information and Library Science*, 34(4), 457–473. <https://doi.org/10.1353/ils.2010.0008>
- Schilling, K., & Applegate, R. (2007). *Evaluating library instruction: Measures for assessing educational quality and impact*. [Conference paper]. The Thirteenth National Conference of the Association of College and Research Libraries, Baltimore, MD, United States. <http://www.ala.org/acrl/conferences/confsandpreconfs/national/baltimore/baltimore>
- Schilling, K., & Applegate, R. (2012). Best methods for evaluating educational impact: A comparison of the efficacy of commonly used measures of library instruction. *Journal of the Medical Library Association*, 100(4), 258–269. <https://doi.org/10.3163/1536-5050.100.4.007>
- SCONUL Working Group on Information Literacy. (2011). *The SCONUL seven pillars of information literacy: Core model for higher education*. SCONUL. <https://www.sconul.ac.uk/sites/default/files/documents/coremodel.pdf>
- Smith, J. K., Given, L. M., Julien, H., Ouellette, D., & DeLong, K. (2013). Information literacy proficiency: Assessing the gap in high school students' readiness for undergraduate academic work. *Library & Information Science Research*, 35(2), 88–96. <https://doi.org/10.1016/j.lisr.2012.12.001>
- Streiner, D. L. (2003). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment*, 80(3), 217–222. [https://doi.org/10.1207/S15327752JPA8003\\_01](https://doi.org/10.1207/S15327752JPA8003_01)
- Streiner, D. L., & Kottner, J. (2014). Recommendations for reporting the results of studies of instrument and scale development and testing. *Journal of Advanced Nursing*, 70(9), 1970–1979. <https://doi.org/10.1111/jan.12402>
- Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, 56(2), 197–208. <https://doi.org/10.1177/0013164496056002001>
- Thompson, S. (2003). *Information literacy meeting of experts: Conference report*. [Conference paper]. Information Literacy Meeting of Experts, Prague, Czech Republic. [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/themes/info\\_lit\\_meeting\\_prague\\_2003.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/themes/info_lit_meeting_prague_2003.pdf)

Thorsen, S. V., & Bjorner, J. B. (2009). Reliability of the Copenhagen Psychosocial Questionnaire. *Scandinavian Journal of Public Health, 38*(3\_suppl), 25–32. <https://doi.org/10.1177/1403494809349859>

Walsh, A. (2009). Information literacy assessment: Where do we start? *Journal of Librarianship and Information Science, 41*(1), 19–28. <https://doi.org/10.1177/0961000608099896>

Walton, G., & Hepworth, M. (2012). Using assignment data to analyse a blended information literacy intervention: A quantitative approach. *Journal of Librarianship and Information Science, 45*(1), 53–63. <https://doi.org/10.1177/0961000611434999>

## Appendix

### Appendix A

#### Exploratory factor analyses

Contents:

1. [Exploratory factor analysis on the IL test items – Pilot sample \(17 items\)](#)
2. [Exploratory factor analysis on the IL test items – Pretest sample \(17 items\)](#)
3. [Exploratory factor analysis on the complete IL test – Undergraduate and graduate samples \(21 items\)](#)
4. [Code for the analyses](#)
5. [References](#)

#### 1. Exploratory factor analysis on the IL test items – Pilot sample (17 items)

Raw data and correlation matrices used in these analyses are available Nierenberg, Låg, and Dahl, 2021.

After excluding items on the basis of low item-total correlations and/or extreme P-values, we performed exploratory factor analyses (EFA) on the 17 remaining items using the data from the Pilot sample. The aim was to explore the dimensionality of the IL test, if it has any, and possibly to further reduce the number of items.

Since items in the IL test are scored dichotomously as correct or incorrect, we used a matrix of tetrachoric correlations among items as the basis for the EFA.

All factor analyses were done using the psych package for R (Revelle, 2020).

##### 1.1 Data suitability for FA

With a sample size of 268 observations, the observation-to-item ratio is 15.76 for these analyses.

We used a number of different criteria (see e.g., Dziuban and Shirkey, 1974) to assess the matrix' suitability for FA. Visual inspection of the matrix revealed few sizeable correlations (only 33 of 136 or 24% were stronger than .3, and only 1 was stronger than .5). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .56, and hardly acceptable. The Bartlett test of sphericity was significant ( $\chi^2(136) = 1389.4, p < .001$ ). Communalities were generally low.

With the exception of a fairly good observation-to-item ratio and a significant Bartlett test (although the latter is not a very reliable indicator), the other indicators seem to cast some doubt on the factorability of these data. We nevertheless decided to perform them, keeping in mind that any sensible solution would need to be cautiously endorsed and preferably replicated.



## 1.2 Extraction and rotation methods

Since items in the IL test are scored dichotomously, multivariate normality is violated. Therefore, the principal axis factoring method was chosen (cf. Fabrigar, Wegener, MacCallum & Strahan, 1999). We had no strong prior assumptions regarding the intercorrelatedness of potential factors, and decided to allow solutions with correlated factors, using an oblique (the Promax with Kaiser normalization) rotation.

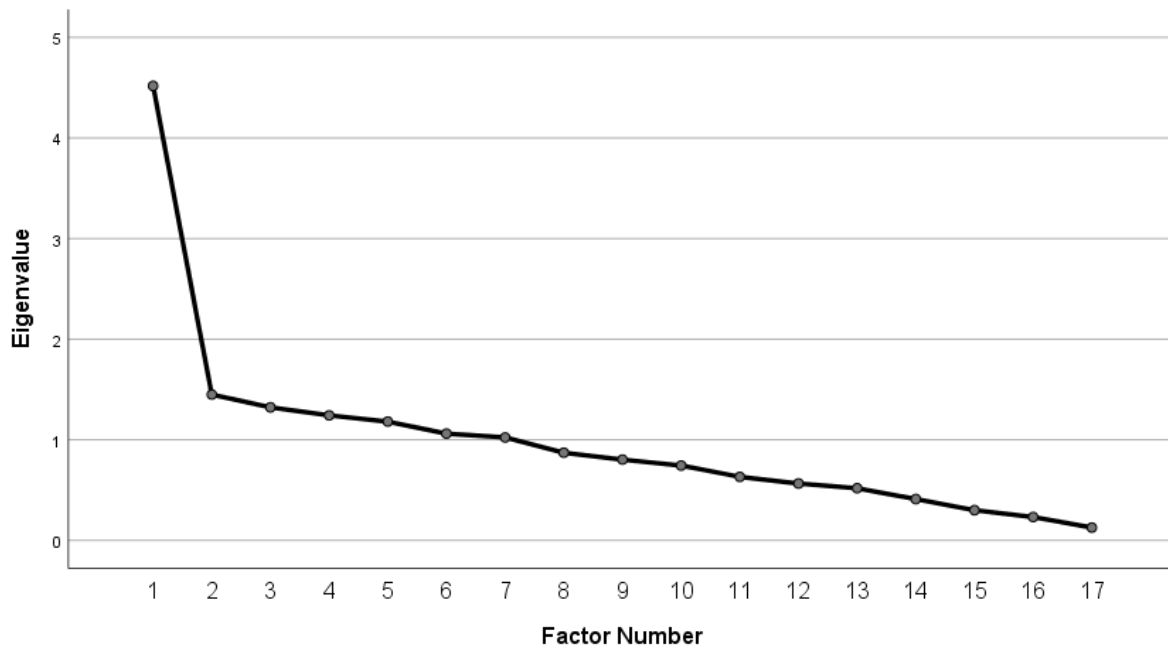
## 1.3 Number of factors

The 17 items remaining after the selection process described in the main manuscript were all related to three core aspects of IL. One might, therefore, expect the most interpretable factor **solution** to be a 3-factor solution. On the other hand, if IL is conceived of as "one underlying variable" or "singular construct" (Hollis, 2018, s. 76-77), a latent attribute that causes the observable behaviours that IL tests measure, then one would expect a strong first factor, with remaining factors accounting for very little variance (Streiner, 2003). A third possibility is that a relatively broadly conceived source oriented IL test measures a number of rather different facets that may or may not be related to each other in any other way than that they tend to be seen as part of IL, in which case we should not expect any clear or meaningful factor structure.

Given that all of these are reasonable a priori possibilities, we hoped in addition to rely on initial eigenvalues in deciding the number of factors to extract. Initial eigenvalues and cumulative percentage of variance explained are reported in Table A1. Figure A1 displays the scree plot.

**Table A1:** Initial eigenvalues and percentage of variance explained for factors extracted from pilot data on 17 items from the IL test

Factor	Eigenvalue	Cumulative % variance
1	4.26	26.6
2	1.41	35.5
3	1.31	43.7
4	1.22	51.3
5	1.13	58.3
6	1.05	64.8



**Figure A1:** Scree plot of initial eigenvalues for factors extracted from pilot data on 17 items from the IL test

The scree plot seems to display only one clear "elbow", indicating that a one-factor solution may be preferable. However, the first factor explains only 26.6% of total variance. And, while subsequent factors increase explained variance only marginally, it still seemed sensible to explore several different solutions.

We explored solutions with everything from 1 to 5 factors. Results for one-, two- and three-factor solutions, which seemed to us marginally more sensible than the other solutions, are reported in the subsections below.

**One-factor solution.** Factor loadings for the one-factor solution are reported in Table A2.

**Table A2:** Factor loadings for a one-factor solution from pilot data on 17 items from the IL test

	Item	Factor 1	IL aspect*	Communalities <sup>+</sup>
14	What kind of information does not need to be cited?	.59	Use	.35
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	.58	Evaluate	.34
5	"The liberal world order will continue..." (... ) Would you say this quote is (... )	.54	Evaluate	.29
6	All of these criteria are useful for assessing the reliability of a journal article, (... )	.52	Evaluate	.27
8	Which reference style should you use in your article if this is not (... )	.52	Use	.27
13	Is it allowed to use an image from a webpage in your article?	.51	Use	.26
10	In which of the sentences below you do not need to cite the source?	.50	Use	.25
3	Which source is the least suitable for finding information for your article on...	.50	Evaluate	.25
4	Which of the following is not necessary in order to critically evaluate a source?	.48	Evaluate	.23
11	John uses paragraphs from his own essay (... ) Is this considered plagiarism?	.47	Use	.22
1	What characterizes a scholarly article?	.47	Evaluate	.22
2	You find these two books (... ). Which of these statements is most correct?	.44	Evaluate	.19
17	You get only two results from (... ) What should you do to get more?	.42	Find	.18
9	What is the most important reason to use sources when writing a paper?	.40	Use	.16
15	You are writing a paper on (... ) Which search gives the best results?	.39	Find	.15
12	[Example shown.] Which source should be cited?	.34	Use	.12
16	The most effective way to follow the newest research on (... )	.16	Find	.03

Notes. Principal axis factor extraction; RMSEA = 0.14 [.13, .15]; TLI = .45

\*These are aspects that the items were originally generated on the basis of

+ Extracted

**Two-factor solution.** Factor loadings for the two-factor solution are reported in Table A3.

**Table A3:** Factor loadings for a 2-factor solution from pilot data on 17 items from the IL test

	Item	Factor 1	Factor 2	IL aspect*	Communalities <sup>†</sup>
14	What kind of information does not need to be cited?	<b>1.16</b>	-.43	Use	.86
13	Is it allowed to use an image from a webpage in your article?	<b>.56</b>	.00	Use	.32
2	You find these two books (...). Which of these statements is most correct?	<b>.55</b>	-.07	Evaluate	.26
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	<b>.40</b>	.12	Use	.24
17	You get only two results from (...) What should you do to get more?	.31	.14	Find	.18
9	What is the most important reason to use sources when writing a paper?	.31	.13	Use	.17
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	-.04	<b>.67</b>	Evaluate	.42
12	[Example shown.] Which source should be cited?	-.23	<b>.60</b>	Use	.23
10	In which of the sentences below do you not need to cite the source?	.02	<b>.52</b>	Use	.29
1	What characterizes a scholarly article?	-.01	<b>.51</b>	Evaluate	.26
3	Which source is the least suitable for finding information for your article on...	.07	<b>.47</b>	Evaluate	.27
4	Which of the following is not necessary in order to critically evaluate a source?	.07	<b>.45</b>	Evaluate	.25
15	You are writing a paper on (...) Which search gives the best results?	-.01	<b>.43</b>	Find	.18
8	Which reference style should you use in your article if this is not (...)	.20	.36	Use	.27
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.26	.32	Evaluate	.28
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.25	.31	Evaluate	.27
16	The most effective way to follow the newest research on (...)	.09	.09	Find	.03

*Notes.* Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; The correlation between the two factors was .68; RMSEA = .14 [.13, .15]; TLI = .46

\* These are aspects that the items were originally generated on the basis of.

† Extracted

**Three-factor solution.** Factor loadings for the three-factor solution are reported in Table A4. Correlations between factors are reported in Table A5.

**Table A4:** Factor loadings for a 3-factor solution from pilot data on 17 items from the IL test

	Item	Factor 1	Factor 2	Factor 3	IL aspect*	Communalities <sup>+</sup>
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	<b>.78</b>	.00	-.22	Evaluate	.51
10	In which of the sentences below you do not need to cite the source?	<b>.67</b>	.08	-.32	Use	.42
12	[Example shown.] Which source should be cited?	<b>.55</b>	-.21	.04	Use	.21
3	Which source is the least suitable for finding information for your article on...	<b>.46</b>	.09	-.02	Evaluate	.27
1	What characterizes a scholarly article?	<b>.45</b>	-.05	.19	Evaluate	.28
4	Which of the following is not necessary in order to critically evaluate a source?	<b>.42</b>	.06	.07	Evaluate	.25
8	Which reference style should you use in your article if this is not (...)	.32	.17	.11	Use	.26
15	You are writing a paper on (...). Which search gives the best results?	.38	-.04	.14	Find	.19
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.27	.22	.12	Evaluate	.27
14	What kind of information does not need to be cited?	-.40	<b>1.24</b>	-.13	Use	.97
13	Is it allowed to use an image from a webpage in your article?	.04	<b>.54</b>	-.04	Use	.31
2	You find these two books (...). Which of these statements is most correct?	-.12	<b>.48</b>	.21	Evaluate	.28
11	John uses paragraphs from his own essay (...). Is this considered plagiarism?	.10	.36	.10	Use	.23
9	What is the most important reason to use sources when writing a paper?	.19	.33	-.13	Use	.19
17	You get only two results from (...). What should you do to get more?	.15	.31	-.01	Find	.18
16	The most effective way to follow the newest research on (...)	-.11	-.06	<b>.61</b>	Find	.30
5	"The liberal world order will continue..." (...). Would you say this quote is (...)	.19	.15	<b>.45</b>	Evaluate	.38

*Notes.* Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; RMSEA = .14 [.13, .15]; TLI = .43

\*These are aspects that the items were originally generated on the basis of.

+Extracted

**Table A5:** Factor correlation matrix

Factor	1	2	3
1		.66	.42
2			.44
3			

### 1.4 Evaluation of factor solutions from pilot sample data

One notable feature of these results is that none of the factors in the two- or three-factor solutions seem to align sensibly to the core IL aspects from which test items were derived, nor indeed to any other meaningful pattern. Neither does there seem to be any support for endorsing a single-factor solution, partly because of the very modest proportion of variance explained. Overall, the items seem to have mostly small correlations among each other, and this is reflected in generally low factor loadings in all the solutions. Since no very clear factor structure is evident, and because we consider broad coverage of the three source-related core IL aspects paramount, we found no basis in these results for further pruning of the test.

## 2. Exploratory factor analysis on the IL test items – Pretest sample (17 items)

As described in the main manuscript, a revised 21-item version (of which 17 items were identical to the those administered to the Pilot sample) of the IL test was administered to a new sample ( $N = 260$ ) of undergraduate students (the Pretest sample).

Although we were unable to make sense of the factor solutions on the Pilot sample data, we were curious as to whether the solutions from the Pilot sample would reproduce in a separate, similar sample. We wanted to remain open-minded as to whether the solutions produced by the analyses on the Pilot sample data were just artefacts of sample idiosyncrasies, or perhaps in fact systematically related to (some unobvious) properties of the test items themselves. If the solutions reproduced in a separate sample, they would perhaps merit a closer look.

At this point we would like to emphasize the exploratory intent behind these analyses. As stated in the main manuscript and in the first section of this supplement, we had few strong reasons to expect a particular factor structure for the IL test to be more likely to emerge than others. Thus, EFA is clearly the most appropriate analytic approach.

However, since the goal of exploratory factor analysis is to infer the likely structure of an instrument when used in a population, it is important to assess whether the same basic factor structure replicates in a separate sample from that population (Osborne, 2014; Osborne & Fitzpatrick, 2012). Factor solutions are notoriously hard to reproduce, even when there is an adequate sample size and a relatively clear factor structure, although both sample size and structure clarity are likely to increase the probability of reproducing the solution.

The analyses presented here were performed on a matrix of tetrachoric correlations (based on the same 17 items analyzed in the Pilot sample).

### 2.1 Data suitability, extraction and rotation methods

With a sample size of 260 observations, the observation-to-item ratio is 15.29 for these analyses.

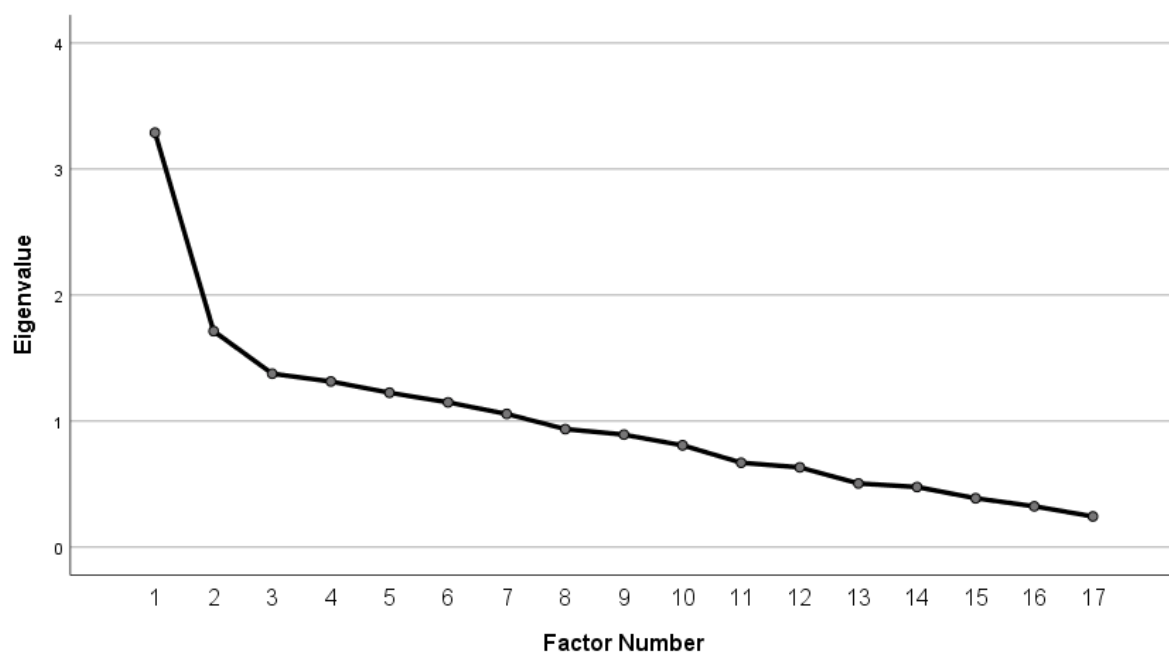
Visual inspection of the matrix again revealed few sizeable correlations (only 19 of 136 or 14% were stronger than .3, and none were stronger than .5). The KMO measure of sampling adequacy was .57, but the Bartlett test of sphericity was significant ( $\chi^2(136) = 869.6, p < .001$ ). Communalities were generally low. Again, the factorability of these data is not very good.

The same general analysis approach (principal axis factoring and Promax rotation with Kaiser normalization) was used to arrive at the factor solutions presented here.

Initial eigenvalues and cumulative percentage of variance explained are reported in Table A6. Figure A2 displays the scree plot.

**Table A6:** Initial eigenvalues and percentage of variance explained for factors extracted from Pretest sample data on 17 items from the IL test

Factor	Eigenvalue	Cumulative % variance
1	3.29	19.3
2	1.71	29.4
3	1.38	37.5
4	1.32	45.3
5	1.23	52.5
6	1.15	49.2



**Figure A2:** Scree plot of initial eigenvalues for factors extracted from the Pretest sample data on 17 items from the IL test

## 2.2 Factor solutions

**One-factor solution.** Factor loadings for the one-factor solution are reported in Table A7.

**Table A7:** Factor loadings for a one-factor solution using data from the Pretest sample on 17 items from the IL test.

	Item	Factor 1	IL aspect*	Communalities*
13	Is it allowed to use an image from a webpage in your article?	.62	Use	.38
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.59	Evaluate	.35
2	You find these two books (...). Which of these statements is most correct?	.54	Evaluate	.29
17	You get only two results from (...) What should you do to get more?	.50	Find	.25
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	.47	Evaluate	.22
3	Which source is the least suitable for finding information for your article on...	.46	Evaluate	.21
14	What kind of information does not need to be cited?	.40	Use	.16
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.35	Evaluate	.12
8	Which reference style should you use in your article if this is not (...)	.35	Use	.12
15	You are writing a paper on (...) Which search gives the best results?	.33	Find	.11
1	What characterizes a scholarly article?	.28	Evaluate	.08
9	What is the most important reason to use sources when writing a paper?	.27	Use	.07
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	.23	Use	.05
12	[Example shown.] Which source should be cited?	.22	Use	.05
16	The most effective way to follow the newest research on (...)	.16	Find	.03
10	In which of the sentences below you do not need to cite the source?	.14	Use	.02
4	Which of the following is not necessary in order to critically evaluate a source?	.12	Evaluate	.01

Notes. Principal axis factor extraction; RMSAE = .11 [.10, .12]; TLI = .39

\*These are aspects that the items were originally generated on the basis of

+ Extracted



**Two-factor solution.** Factor loadings for the two-factor solution are reported in Table A8.

**Table A8:** Factor loadings (pattern matrix) for a 2-factor solution from Pretest sample data on 17 items from the IL test

	Item	Factor 1	Factor 2	IL aspect*	Communalities <sup>+</sup>
2	You find these two books (...). Which of these statements is most correct?	<b>.65</b>	-.02	Evaluate	.41
13	Is it allowed to use an image from a webpage in your article?	<b>.62</b>	.10	Use	.44
3	Which source is the least suitable for finding information for your article on...	<b>.53</b>	-.00	Evaluate	.28
15	You are writing a paper on (...) Which search gives the best results?	<b>.41</b>	-.04	Find	.16
17	You get only two results from (...) What should you do to get more?	<b>.41</b>	.18	Find	.25
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	.36	-.11	Use	.11
14	What kind of information does not need to be cited?	.35	.12	Use	.17
16	The most effective way to follow the newest research on (...)	.21	-.03	Find	.04
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	.08	<b>.55</b>	Evaluate	.34
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.00	<b>.48</b>	Evaluate	.23
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.29	<b>.44</b>	Evaluate	.37
4	Which of the following is not necessary in order to critically evaluate a source?	-.23	<b>.43</b>	Evaluate	.17
12	[Example shown.] Which source should be cited?	-.07	.39	Use	.14
8	Which reference style should you use in your article if this is not (...)	.09	.36	Use	.16
1	What characterizes a scholarly article?	.02	.34	Evaluate	.12
9	What is the most important reason to use sources when writing a paper?	.09	.24	Use	.08
10	In which of the sentences below you do not need to cite the source?	-.03	.22	Use	.05

Notes. Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; The correlation between the two factors was .36; RMSEA = .11 [.10, .12]; TLI = .43

\* These are aspects that the items were originally generated on the basis of.

<sup>+</sup> Extracted

**Three-factor solution.** Factor loadings for the three-factor solution are reported in Table A9. Correlations between factors are reported in Table A10.

**Table A9:** Factor loadings (pattern matrix) for a 3-factor solution from Pretest sample data on 17 items from the IL test

	Item	Factor 1	Factor 2	Factor 3	IL aspect*	Communalities <sup>+</sup>
2	You find these two books (...). Which of these statements is most correct?	<b>.80</b>	-.17	-.03	Evaluate	.56
3	Which source is the least suitable for finding information for your article on...	<b>.52</b>	-.01	.02	Evaluate	.26
13	Is it allowed to use an image from a webpage in your article?	<b>.49</b>	.21	.26	Use	.41
14	What kind of information does not need to be cited?	<b>.40</b>	.06	-.01	Use	.18
17	You get only two results from (...) What should you do to get more?	<b>.40</b>	.17	.04	Find	.24
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	.34	-.10	.12	Use	.12
16	The most effective way to follow the newest research on (...)	.23	-.05	-.02	Find	.05
4	Which of the following is not necessary in order to critically evaluate a source?	-.31	<b>.53</b>	.07	Evaluate	.24
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	.13	<b>.49</b>	-.06	Evaluate	.32
12	[Example shown.] Which source should be cited?	-.10	<b>.42</b>	-.01	Use	.16
8	Which reference style should you use in your article if this is not (...)	.04	<b>.42</b>	.11	Use	.19
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.32	.39	-.06	Evaluate	.37
9	What is the most important reason to use sources when writing a paper?	.06	.28	.05	Use	.09
10	In which of the sentences below you do not need to cite the source?	-.04	.24	.04	Use	.05
1	What characterizes a scholarly article?	.12	.24	-.15	Evaluate	.13
15	You are writing a paper on (...) Which search gives the best results?	.24	.26	<b>.97</b>	Find	1.01
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.12	.35	-.34	Evaluate	.36

Notes. Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; RMSEA = .10 [.09, .12]; TLI = .48

\*These are aspects that the items were originally generated on the basis of.

\*Extracted

**Table A10:** Factor correlation matrix

Factor	1	2	3
1		.40	-.05
2			-.14
3			

### 2.3 Evaluation of factor solutions from Pretest sample data and comparison to solutions from Pilot sample data

As with the pilot sample data, there is no readily interpretable factor solution, which may simply be due to the fact that the correlation matrix is not very suitable for factorization. Comparing the factor solutions from the two samples, one would be hard pressed to say that the few similarities reach a minimum replicability threshold of the same basic factor structure. For instance, comparing the two-factor solutions, several items load on different factors (e.g., item 15 and 3 are grouped with items 2, 14, 13, and 11 in the Pretest sample solution, but not in the Pilot sample solution). Even when items group on the same factor in solutions from the two samples, loading magnitudes are somewhat different.

In sum, when considering explored solutions from the Pilot sample data and the Pretest sample data, there seem to be no readily interpretable solutions, and the solutions emerging from the two samples do not match very well. No clear dimensionality is evident for the IL test based on these analyses.

## 3. Exploratory factor analysis on the complete IL test – Undergraduate and graduate samples (21 items)

In spite of the fact that the observation-to-item ratios of the analyses described in the previous section are fairly good compared to what seems to be common practice in fields that employ factor analysis (cf. for instance Ford, MacCallum & Tait, 1986, who found that 56% of analyses published over a 10-year period had observation-to-item ratios of less than 10:1), they are still slightly short of ideals recommended by some (e.g., the 500 observation limit for considering a sample size "very good" according to Comrey & Lee, 1992, or the 400 observation minimum suggested by Aleamoni, 1976).

Furthermore, although no readily interpretable structure emerged from the analyses described in the previous sections, we deemed it possible that one might emerge from i) a larger sample; ii) a sample including more experienced students likely to manifest higher levels of IL; iii) a matrix based on the full 21-item test, including revised items.

We therefore merged the Pretest sample ( $N = 260$ ) with the Graduate student sample ( $N = 366$ ) and again explored the possible dimensionality of the IL test using EFA.

### 3.1 Data suitability for FA

With a total sample size of 626 observations, the observation-to-item ratio is 29.81 for these analyses.

Visual inspection of the matrix revealed few sizeable correlations (only 39 of 210 or 18.6% were stronger than .3, and only 3 of 210 or 1.4% were stronger than .5). The KMO measure of sampling adequacy was .76, which is typically considered adequate. The Bartlett test of sphericity was significant ( $\chi^2(210) = 3668.4, p < .001$ ). Communalities, however, were generally low.

Again, despite a conventionally adequate KMO-value, the limited number of intercorrelations among the items indicates relatively poor factorability. We nevertheless decided to perform the

analyses, again keeping in mind that any sensible solution would need to be cautiously endorsed and preferably replicated.

### 3.2 Extraction and rotation methods

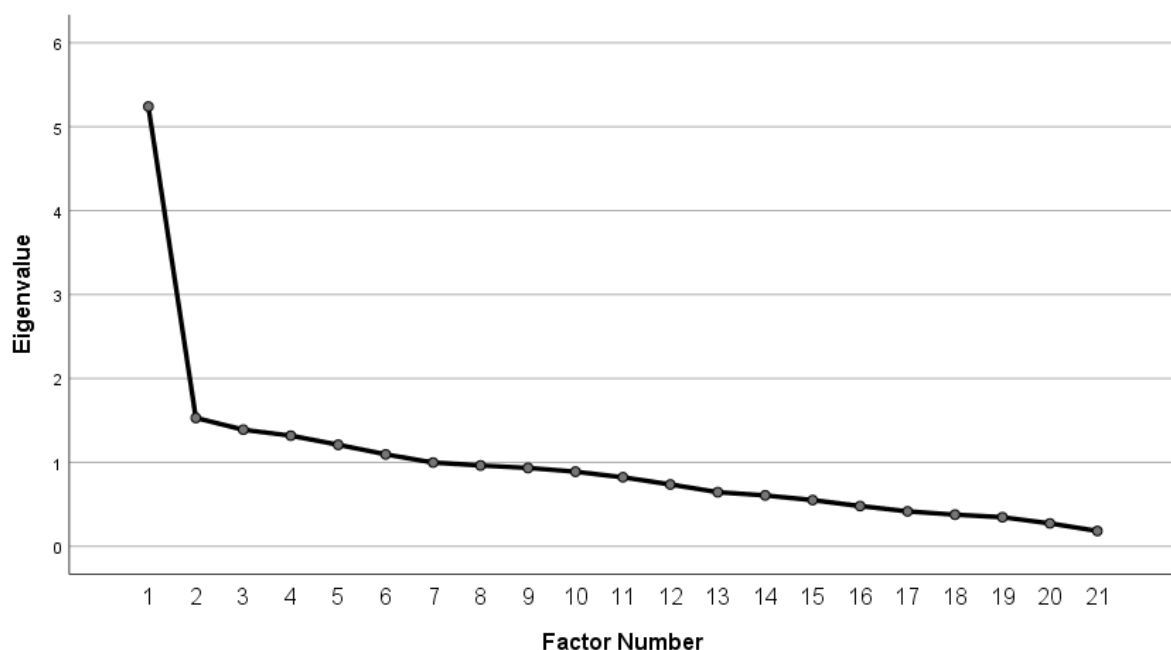
The considerations regarding extraction and rotation methods from the analyses described above were applied to the current sample.

### 3.3 Number of factors

Again, given the varied reasonable a priori possibilities outlined in the first section of this supplement, we hoped in addition to rely on initial eigenvalues in deciding the number of factors to extract. Initial eigenvalues and cumulative percentage of variance explained are reported in Table A11. Figure A3 displays the scree plot.

**Table A11:** Initial eigenvalues and percentage of variance explained for factors extracted from data on the complete 21 item IL test.

Factor	Eigenvalue	Cumulative % variance
1	5.25	25.0
2	1.53	32.2
3	1.39	38.9
4	1.32	45.1
5	1.21	50.9
6	1.10	56.1



**Figure A3:** Scree plot of initial eigenvalues for factors extracted from data on the complete 21-item IL test

The scree plot seems to display only one clear "elbow", indicating that a one-factor solution may be preferable. However, the first factor explains only 25% of the total variance. And, while subsequent factors increase explained variance only marginally, it still seemed sensible to explore several different solutions.

We explored solutions with everything from 1 to 5 factors. Results for one-, two- and three-factor solutions are reported in the subsections below.

**One-factor solution.** Factor loadings for the one-factor solution are presented in Table A12.

**Table A12:** Factor loadings for a one-factor solution using data from the complete 21-item IL test

	Item	Factor 1	IL aspect*	Communalities <sup>+</sup>
13	Is it allowed to use an image from a webpage in your article?	.72	Use	.52
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.67	Evaluate	.45
8	Which reference style should you use in your article if this is not (...)	.63	Use	.39
20	Which statement is correct? [Statements about searching.]	.62	Find	.39
21	(...) If you do not get sufficient results the first time, what should you do?	.61	Find	.37
15	You are writing a paper on (...) Which search gives the best results?	.60	Find	.36
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	.55	Evaluate	.30
1	What characterizes a scholarly article?	.48	Evaluate	.23
3	Which source is the least suitable for finding information for your article on...	.45	Evaluate	.20
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.44	Evaluate	.20
17	You get only two results from (...) What should you do to get more?	.41	Find	.17
4	Which of the following is not necessary in order to critically evaluate a source?	.40	Evaluate	.16
10	In which of the sentences below you do not need to cite the source?	.37	Use	.14
16	The most effective way to follow the newest research on (...)	.36	Find	.13
2	You find these two books (...). Which of these statements is most correct?	.34	Evaluate	.12
18	If you wrote comp* in the search field of a database such as (...)	.32	Find	.10
14	What kind of information does not need to be cited?	.31	Use	.10
9	What is the most important reason to use sources when writing a paper?	.29	Use	.09
19	In which case is it not recommended to use OR between words (...)	.26	Find	.07
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	.22	Use	.05
12	[Example shown.] Which source should be cited?	.14	Use	.02

Notes. Principal axis factor extraction

\*These are aspects that the items were originally generated on the basis of; RMSEA = .11 [.10, .11]; TLI = .57

\*Extracted

**Two-factor solution.** Factor loadings for the two-factor solution are reported in Table A13.

**Table A13:** Factor loadings (pattern matrix) for a two-factor solution using data from the complete 21-item IL test

	Item	Factor 1	Factor2	IL aspect*	Communalities <sup>+</sup>
18	If you wrote comp* in the search field of a database such as (...)	<b>.83</b>	-.50	Find	.41
8	Which reference style should you use in your article if this is not (...)	<b>.67</b>	.01	Use	.46
15	You are writing a paper on (...) Which search gives the best results?	<b>.55</b>	.10	Find	.38
4	Which of the following is not necessary in order to critically evaluate a source?	<b>.48</b>	-.06	Evaluate	.20
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	<b>.45</b>	.15	Evaluate	.30
21	(...) If you do not get sufficient results the first time, what should you do?	<b>.44</b>	.22	Find	.37
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	<b>.42</b>	.31	Evaluate	.45
10	In which of the sentences below you do not need to cite the source?	.39	.01	Use	.15
16	The most effective way to follow the newest research on (...)	.37	.02	Find	.15
1	What characterizes a scholarly article?	.35	.17	Evaluate	.23
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.32	.16	Evaluate	.19
19	In which case is it not recommended to use OR between words (...)	.22	.06	Find	.07
12	[Example shown.] Which source should be cited?	.11	.05	Use	.02
3	Which source is the least suitable for finding information for your article on...	-.27	<b>.83</b>	Evaluate	.48
13	Is it allowed to use an image from a webpage in your article?	.21	<b>.61</b>	Use	.58
20	Which statement is correct? [Statements about searching.]	.20	<b>.49</b>	Find	.42
2	You find these two books (...). Which of these statements is most correct?	-.03	<b>.42</b>	Evaluate	.17
14	What kind of information does not need to be cited?	.01	.35	Use	.13
11	John uses paragraphs from his own essay (...) Is this considered plagiarism?	-.04	.29	Use	.07
17	You get only two results from (...) What should you do to get more?	.19	.27	Find	.17
9	What is the most important reason to use sources when writing a paper?	.11	.21	Use	.09

Notes. Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; The correlation between the two factors was .58; RMSEA = .10 [.10, .11]; TLI = .59

\* These are aspects that the items were originally generated on the basis of.

<sup>+</sup> Extracted

**Three-factor solution.** Factor loadings for the three-factor solution are reported in Table A14. Correlations between factors are reported in Table A15.

**Table A14:** Factor loadings (pattern matrix) for a 3-factor solution using data from the complete 21-item IL test

	Item	Factor 1	Factor 2	Factor 3	IL aspect*	Communalities <sup>+</sup>
1 3	Is it allowed to use an image from a webpage in your article?	<b>.82</b>	-.11	.21	Use	.75
2 0	Which statement is correct? [Statements about searching.]	<b>.70</b>	-.08	.20	Find	.51
3	Which source is the least suitable for finding information for your article on...	<b>.51</b>	.42	-.39	Evaluate	.50
6	All of these criteria are useful for assessing the reliability of a journal article, (...)	.32	.22	.28	Evaluate	.44
2 1	(...) If you do not get sufficient results the first time, what should you do?	.25	.19	.30	Find	.37
1 1	John uses paragraphs from his own essay (...) Is this considered plagiarism?	.34	-.08	.00	Use	.09
2	You find these two books (...). Which of these statements is most correct?	.29	.21	-.10	Evaluate	.16
1	What characterizes a scholarly article?	-.06	<b>.53</b>	.10	Evaluate	.30
7	You are writing a paper about Arctic ice melting. Who most likely has the least...	-.02	<b>.47</b>	.21	Evaluate	.35
4	Which of the following is not necessary in order to critically evaluate a source?	-.21	<b>.45</b>	.26	Evaluate	.27
1 4	What kind of information does not need to be cited?	.10	<b>.42</b>	-.16	Use	.18
1 7	You get only two results from (...) What should you do to get more?	.14	.30	.05	Find	.18
9	What is the most important reason to use sources when writing a paper?	.01	.38	-.05	Use	.13
5	"The liberal world order will continue..." (...) Would you say this quote is (...)	.06	.31	.16	Evaluate	.21
1 9	In which case is it not recommended to use OR between words (...)	-.12	.39	.04	Find	.12
1 2	[Example shown.] Which source should be cited?	.05	.05	.07	Use	.02
1 8	If you wrote comp* in the search field of a database such as (...)	-.22	-.04	<b>.71</b>	Find	.40
1 5	You are writing a paper on (...) Which search gives the best results?	.29	-.02	<b>.49</b>	Find	.44
8	Which reference style should you use in your article if this is not (...)	.01	.33	<b>.44</b>	Use	.46
1 6	The most effective way to follow the newest research on (...)	.24	-.16	.39	Find	.21
1 0	In which of the sentences below you do not need to cite the source?	.06	.11	.28	Use	.15

Notes. Principal axis factor extraction; Promax rotation with Kaiser normalization; Factor loadings larger than .40 are in bold; RMSEA = .10 [.09, .11]; TLI = .62

\* These are aspects that the items were originally generated on the basis of.

+ Extracted



**Table A15:** Factor correlation matrix

Factor	1	2	3
1		.59	.44
2			.51
3			

### 3.4 Evaluation of factor solutions from combined samples (pretest sample and graduate sample).

Despite substantially increased sample size, the factorability of the correlations among the IL test items remains poor. Again, a notable feature of the results is that none of the factors in the two- or three-factor solutions seem to align sensibly to the core IL aspects from which test items were derived. And again there seems to be no support for endorsing a single-factor solution, due to the very low proportion of variance explained. Items in the IL test have few and small correlations amongst each other, and this is reflected in generally low factor loadings in all the solutions.

## 4. Code for the analyses reported above

```
### Exploratory factor analyses of IL-test items ###
```

```
library(psych)
library(psychTools)
library(GPArotation)
```

```
### Analyses of data from the pilot sample (17 items) ###
```

```
pilot_data <- read.file() #Load the file ILpilotdata_forEFA_17items.txt
tet_mat <- tetrachoric(pilot_data) #For visual inspection, scree plot generation and sampling
adequacy assessment
KMO(tet_mat$rho) #Calculate the KMO measure of sampling adequacy
cortest.bartlett(tet_mat$rho, n = 268) #Calculate the Bartlett test of sphericity
scree(tet_mat$rho, factors=FALSE, pc=TRUE) #Generating scree plot from the correlation
matrix
```

```
## Factor analysis - One-factor solution ##
```

```
fa_pilot_17_1f <- fa(pilot_data, nfactors=1, rotate="promax", scores="regression", fm="pa",
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation
with Kaiser normalization.
fa_pilot_17_1f #Display factor analysis output
fa_pilot_17_1f$e.values #Printing eigenvalues
```

```
## Factor analysis - Two-factor solution ##
```

```
fa_pilot_17_2f <- fa(pilot_data, nfactors=2, rotate="promax", scores="regression", fm="pa",
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation
with Kaiser normalization.
fa_pilot_17_2f #Display factor analysis output
fa_pilot_17_2f$e.values #Printing eigenvalues
```

```
## Factor analysis - Three-factor solution ##
```

```
fa_pilot_17_3f <- fa(pilot_data, nfactors=3, rotate="promax", scores="regression", fm="pa",
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation
with Kaiser normalization.
fa_pilot_17_3f #Display factor analysis output
```

```
fa_pilot_17_3f$e.values #Printing eigenvalues
```

```
####Analyses of data from the undergraduate (pretest) sample (17 items) ###  
ugrad_data <- read.file() #Load the file Undergrad_forEFA_17items.txt  
tet_mat <- tetrachoric(ugrad_data) #For visual inspection, scree plot generation and sampling  
adequacy assessment  
KMO(tet_mat$rho) #Calculate the KMO measure of sampling adequacy  
cortest.bartlett(tet_mat$rho, n = 260) #Calculate the Bartlett test of sphericity  
scree(tet_mat$rho, factors=FALSE, pc=TRUE) #Generating scree plot from the correlation  
matrix
```

```
## Factor analysis - One-factor solution ##
```

```
fa_ugrad_17_1f <- fa(ugrad_data, nfactors=1, rotate="promax", scores="regression", fm="pa",  
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation  
with Kaiser normalization.
```

```
fa_ugrad_17_1f #Display factor analysis output
```

```
fa_ugrad_17_1f$e.values #Printing eigenvalues
```

```
## Factor analysis - Two-factor solution ##
```

```
fa_ugrad_17_2f <- fa(ugrad_data, nfactors=2, rotate="promax", scores="regression", fm="pa",  
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation  
with Kaiser normalization.
```

```
fa_ugrad_17_2f #Display factor analysis output
```

```
fa_ugrad_17_2f$e.values #Printing eigenvalues
```

```
## Factor analysis - Three-factor solution ##
```

```
fa_ugrad_17_3f <- fa(ugrad_data, nfactors=3, rotate="promax", scores="regression", fm="pa",  
cor="tet") #Calling the factor analysis function with principal axis factoring and promax rotation  
with Kaiser normalization.
```

```
fa_ugrad_17_3f #Display factor analysis output
```

```
fa_ugrad_17_3f$e.values #Printing eigenvalues
```

```
####Analyses of data from the merged (undergrad pretest + graduate) sample (21 items) ###
```

```
merged_data <- read.file() #Load the file Merged_forEFA_21items.txt  
tet_mat <- tetrachoric(merged_data) #For visual inspection, scree plot generation and sampling  
adequacy assessment
```

```
KMO(tet_mat$rho) #Calculate the KMO measure of sampling adequacy
```

```
cortest.bartlett(tet_mat$rho, n = 626) #Calculate the Bartlett test of sphericity
```

```
scree(tet_mat$rho, factors=FALSE, pc=TRUE) #Generating scree plot from the correlation  
matrix
```

```
## Factor analysis - One-factor solution ##
```

```
fa_merged_21_1f <- fa(merged_data, nfactors=1, rotate="promax", scores="regression",  
fm="pa", cor="tet") #Calling the factor analysis function with principal axis factoring and promax  
rotation with Kaiser normalization.
```

```
fa_merged_21_1f #Display factor analysis output
```

```
fa_merged_21_1f$e.values #Printing eigenvalues
```

```
## Factor analysis - Two-factor solution ##
```

```
fa_merged_21_2f <- fa(merged_data, nfactors=2, rotate="promax", scores="regression",  
fm="pa", cor="tet") #Calling the factor analysis function with principal axis factoring and promax  
rotation with Kaiser normalization.
```

```
fa_merged_21_2f #Display factor analysis output
```

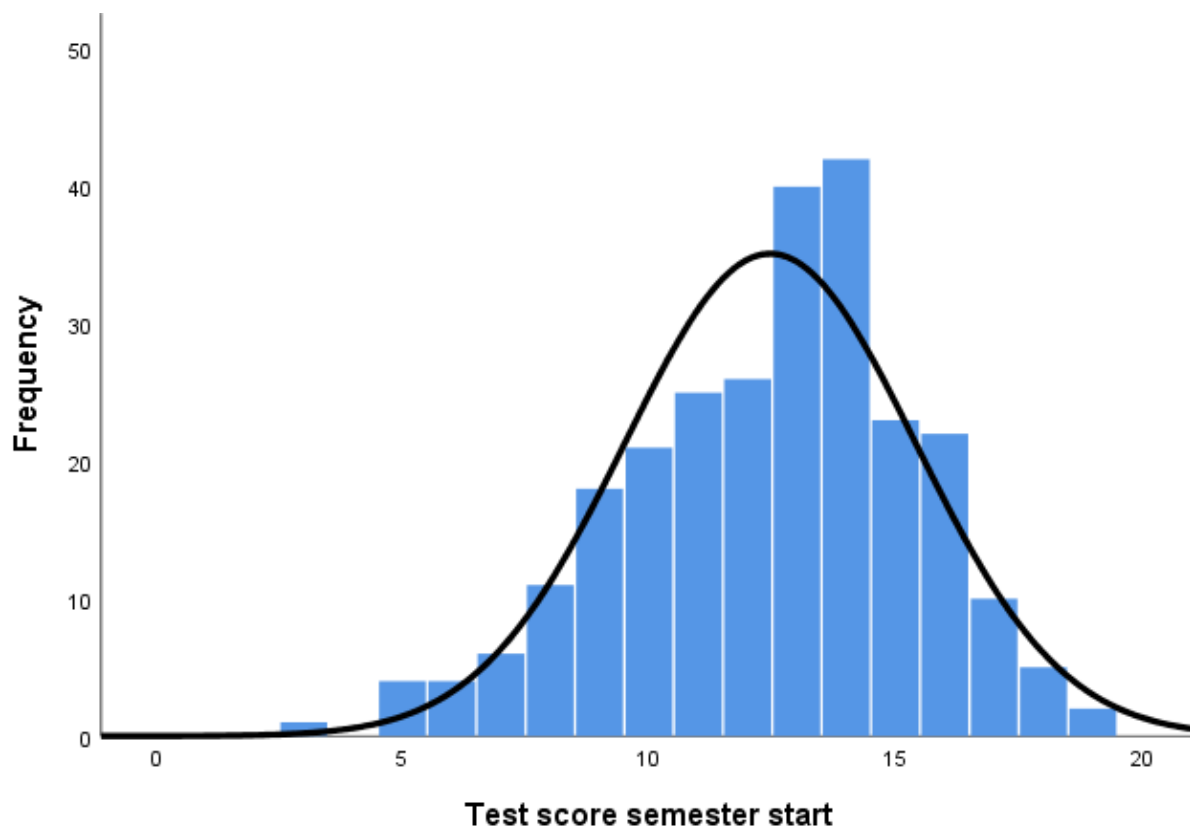
```
fa_merged_21_2f$e.values #Printing eigenvalues
```

```
## Factor analysis - Three-factor solution ##
fa_merged_21_3f <- fa(merged_data, nfactors=3, rotate="promax", scores="regression",
fm="pa", cor="tet") #Calling the factor analysis function with principal axis factoring and promax
rotation with Kaiser normalization.
fa_merged_21_3f #Display factor analysis output
fa_merged_21_3f$e.values #Printing eigenvalues
```

## 5. References

- Aleamoni, L. M. (1976). The relation of sample size to the number of variables in using factor analysis techniques. *Educational and Psychological Measurement*, 36, 879–883.  
<https://doi.org/10.1177/001316447603600410>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum Associates.
- Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 81, 358–361.  
<https://doi.org/10.1037/h0036316>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.  
<https://doi.org/10.1037/1082-989X.4.3.272>
- Ford, J. K., Maccallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291–314.  
<https://doi.org/10.1111/j.1744-6570.1986.tb00583.x>
- Hollis, H. (2018). Information literacy as a measurable construct. *Journal of Information Literacy*, 12(2), 76-88. <https://doi.org/10.11645/12.2.2409>
- Nierenberg, E., Låg, T., & Dahl, T. I. (2021). *Replication Data for: Knowing and doing: The development of information literacy measures to assess knowledge and practice* (DataverseNO; Version V2) [Data set, survey and evaluation measures].  
<https://doi.org/10.18710/L60VDI>
- Osborne, J. W. (2014). *Best practices in exploratory factor analysis* [Kindle edition].  
<https://amazon.com>
- Osborne, J. W., & Fitzpatrick, D. C. (2012). Replication analysis in exploratory factor analysis: What it is and why it makes your analysis better. *Practical Assessment, Research & Evaluation*, 17(15), 1-8. <https://pareonline.net/getvn.asp?v=17%26n=15>
- Revelle, W. (2020). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych>  
Version = 2.0.8
- Streiner, D. L. (2003). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment*, 80, 217–222.  
[https://doi.org/10.1207/S15327752JPA8003\\_01](https://doi.org/10.1207/S15327752JPA8003_01)

## Appendix B



**Figure B1:** Distribution of undergraduate test scores at semester start

$N = 260$

Mean = 12.46

Median = 13

Mode = 14

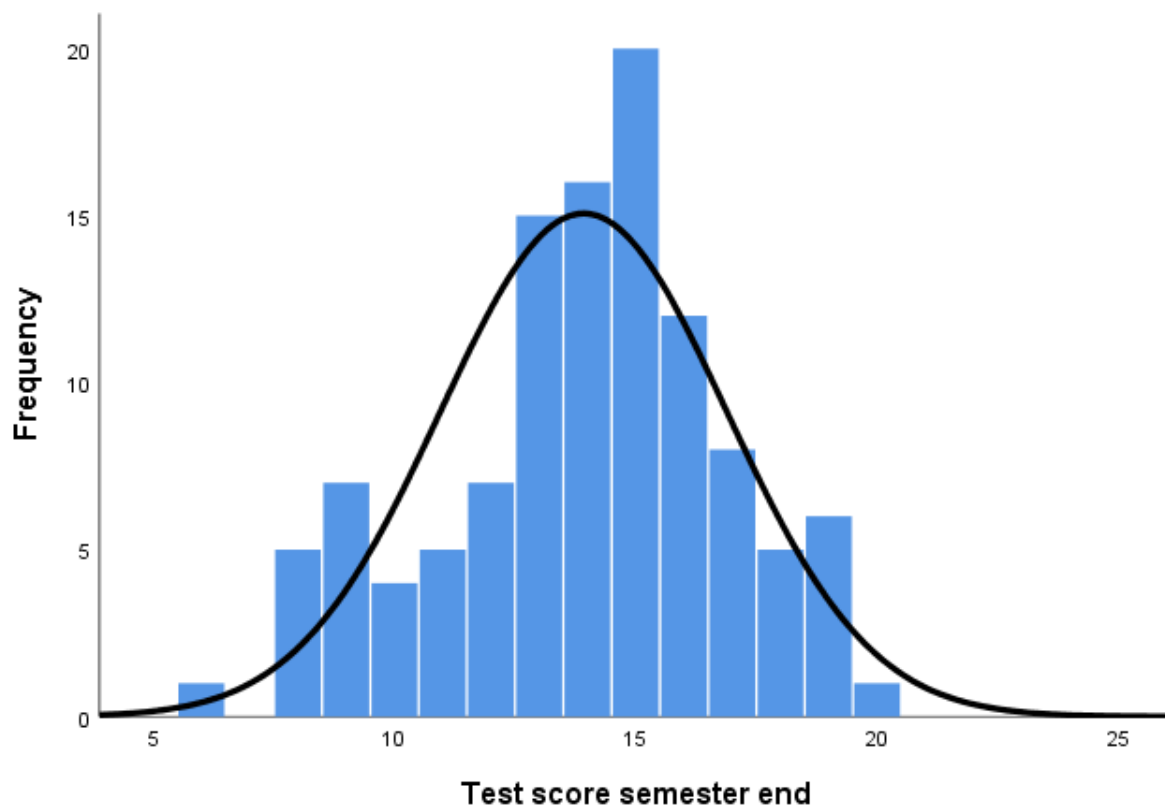
Std. Deviation = 2.953

Skewness = -.424

Std. Error of Skewness = .151

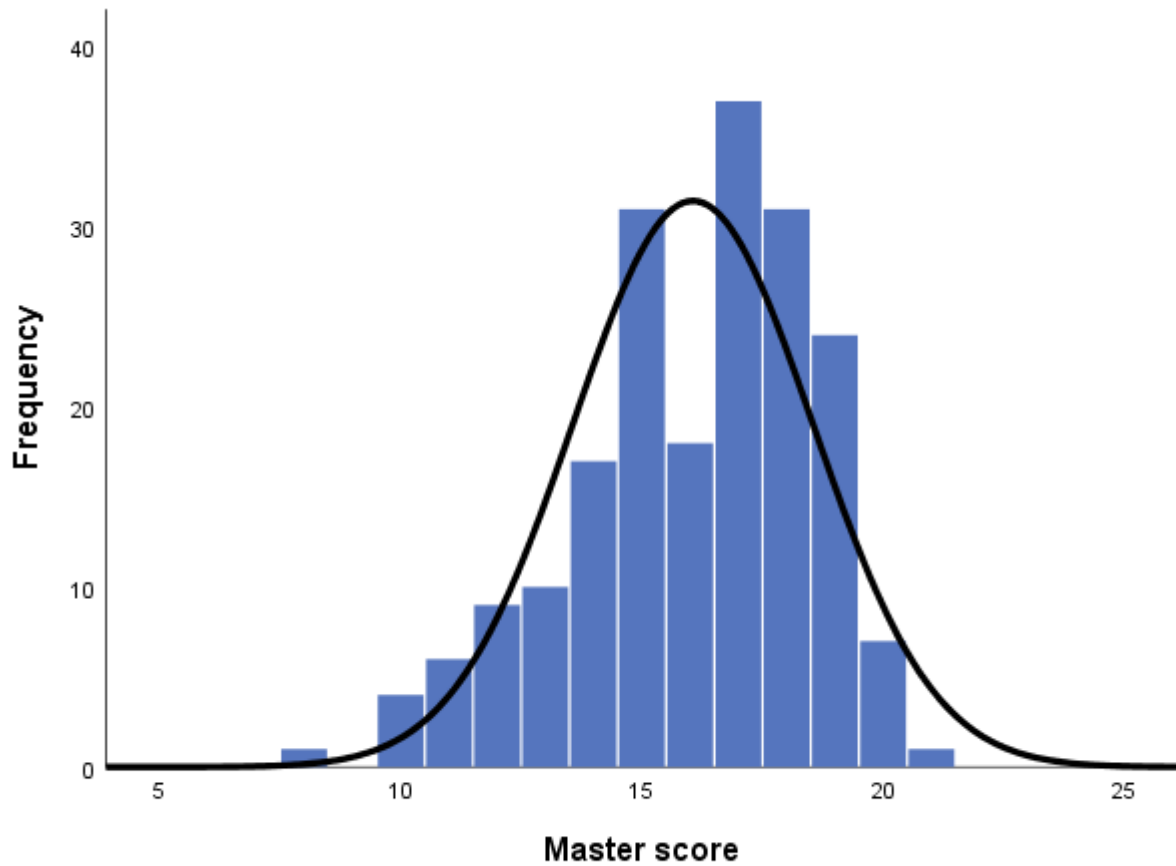
Kurtosis = -.034

Std. Error of Kurtosis = .301



**Figure B2:** Distribution of undergraduate test scores at semester end

N = 112  
 Mean = 13.92  
 Median = 14  
 Mode = 15  
 Std. Deviation = 2.966  
 Skewness = -.378  
 Std. Error of Skewness = .228  
 Kurtosis = -.217  
 Std. Error of Kurtosis = .453



**Figure B3:** Distribution of Master's student scores

$N = 196$

Mean = 16.06

Median = 17

Mode = 17

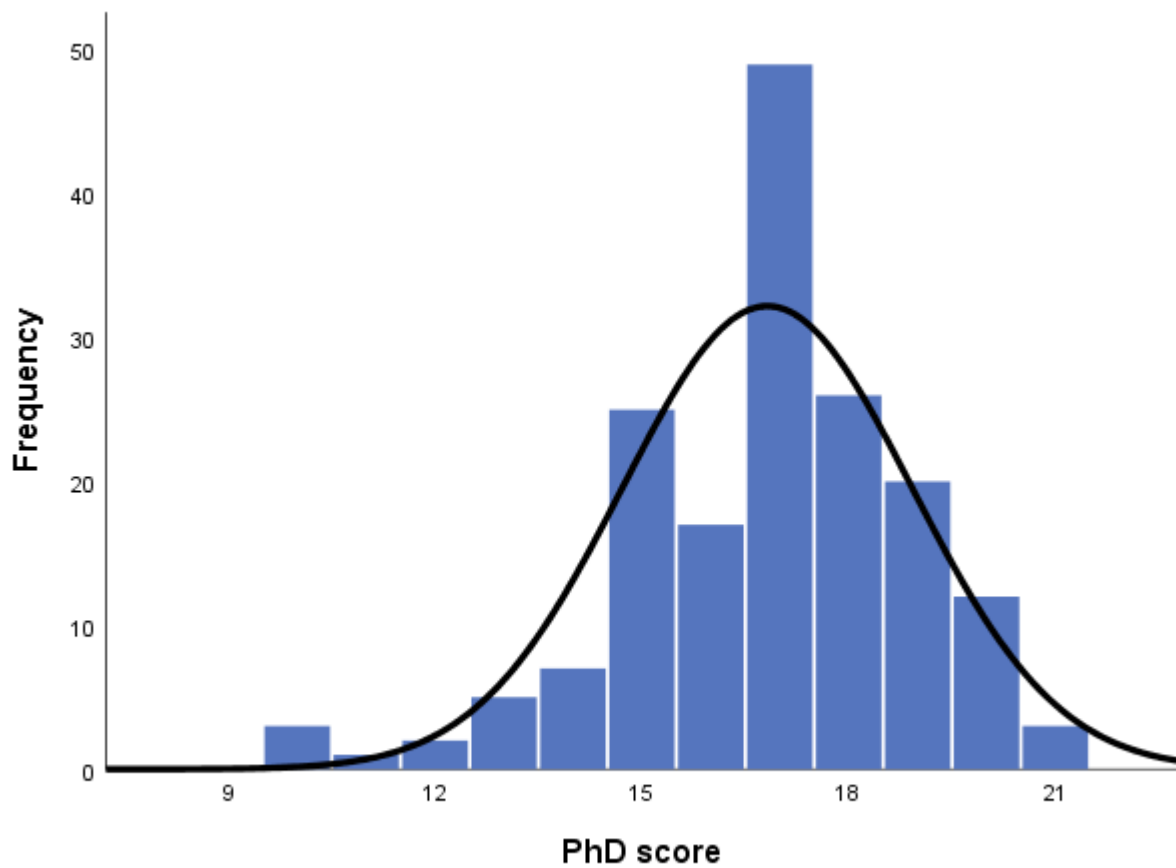
Std. Deviation = 2.485

Skewness = -.620

Std. Error of Skewness = .174

Kurtosis = -.034

Std. Error of Kurtosis = .346



**Figure B4:** Distribution of PhD student scores

$N = 170$

Mean = 16.82

Median = 17

Mode = 17

Std. Deviation = 2.106

Skewness = -.711

Std. Error of Skewness = .186

Kurtosis = 1.044

Std. Error of Kurtosis = .370

## Appendix C

### [1. Correlations](#)

### [2. Regressions](#)

#### 1. Correlations

Correlations for seven study variables are reported in Table C1:

- KNOW
  - Information literacy knowledge test, semester-start score (KnowStart)
- DO
  - Source Evaluation total score (SourceEval Total Score)
  - Source Evaluation quality score (SourceEval Quality)
  - Source Evaluation variety score (SourceEval Variety)
  - Source Evaluation frequency score (SourceEval Frequency)
  - Source Use score (SourceUse Score)
- EXPERIENCE
  - Number of semesters of previous higher education (Semesters HE)

**Table C1:** Correlations for study variables, in Pearson's *r*

Variable	<i>n</i>	1	2	3	4	5	6	7
1. KnowStart	260	–						
2. SourceEval Total Score	93	.195	–					
3. SourceEval Quality	93	.272**	.466**	–				
4. SourceEval Variety	93	.150	.932**	.359**	–			
5. SourceEval Frequency	93	.184	.940**	.300**	.943**	–		
6. SourceUse Score	87	.308**	.197	.157	.141	.166	–	
7. Semesters HE	260	.150*	-.033	.022	-.032	-.052	-.105	–

Notes:

\**p* < .05

\*\**p* < .01



### 1.1 Comments on correlations

Aside from correlations between the different source evaluation scores, there are three significant correlations between the analyzed variables:

- The source evaluation quality score and the source use score are positively correlated with the Information literacy (IL) knowledge at semester-start scores. These correlations are statistically significantly, but not strong.
- There is a positive correlation between the Number of semesters of previous higher education and IL knowledge at semester-start scores. This correlation is statistically significant, but weak.

## 2. Regressions

In order to follow up on the associations reported in the correlation matrix, and specifically, to attempt to control for higher education experience, we performed two multiple regressions using IL knowledge test scores and number of semesters of higher education as predictors, and the source evaluation quality score and source use scores as dependent variables.

For both of these analyses there were no indications of multicollinearity or influential cases, errors are independent, and there was no obvious heteroscedasticity. Residuals showed some deviation from normality, particularly for the analysis reported in Table C2.

**Table C2:** Regression coefficients of IL knowledge test scores on Source Evaluation quality score

Variable	Model 1			Model 2		
	<i>B</i>	$\beta$	<i>SE</i>	<i>B</i>	$\beta$	<i>SE</i>
Constant	1.71**		.34	1.71**		.34
IL knowledge semester start score	.07**	.27	.03	.07**	.27	.03
Numbers of semesters in HE				.00	.00	.08
$R^2$	.07			.07		
$\Delta R^2$				.00		

Note.  $N = 93$ . In Model 1, we entered the IL test scores. In Model 2, we entered Number of semesters of higher education as a control variable.

\*\* $p < .01$ .

**Table C3:** Regression coefficients of IL knowledge test scores on Source use score

Variable	Model 1			Model 2		
	<i>B</i>	$\beta$	<i>SE</i>	<i>B</i>	$\beta$	<i>SE</i>
Constant	1.07**		.56	1.73**		.55
IL knowledge semester start score	.13**	.31	.04	.13**	.32	.04
Numbers of semesters in HE				-.17		.14
<i>R</i> <sup>2</sup>	.10			.11		
$\Delta R^2$	.02					

Note. *N* = 87. In Model 1, we entered the IL test scores. In Model 2, we entered Number of semesters of higher education as a control variable.

\*\**p* < .01.

### 2.1 Comments on regressions

None of these models account for much of the variance in the dependent variables (scores on do-measures). Nevertheless, as an illustrative example, and assuming the models generalize, a standardized coefficient of just above .3 between knowledge test at semester start scores and source use scores, allows us to predict almost a third of a standard deviation increase in the ability to use sources for every increase of one standard deviation in IL knowledge. In other words, an increase of just under three points on the knowledge test translates to a 0.34 point increase on the source use score (which has a maximum score of five). While not impressive, it is not entirely trivial.

Note that the attempted control for number of semesters is weak, due to the restricted range of this variable in this sample.

## Paper 2



# Here's the TRIQ: The Tromsø Interest Questionnaire Based on the Four-Phase Model of Interest Development

Tove I. Dahl<sup>\*1</sup> and Ellen Nierenberg<sup>2</sup>

<sup>1</sup>Department of Psychology, UiT The Arctic University of Norway, Tromsø, Norway, <sup>2</sup>Department of Psychology and University Library, UiT The Arctic University of Norway, Tromsø, Norway

## OPEN ACCESS

### Edited by:

Ramón Chacón-Cuberos,  
University of Granada, Spain

### Reviewed by:

Michael Schurig,  
Technical University Dortmund,  
Germany

Mariana Mármol,  
University of Granada, Spain

### \*Correspondence:

Tove I. Dahl  
tove.dahl@uit.no

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

**Received:** 28 May 2021

**Accepted:** 21 September 2021

**Published:** 08 November 2021

### Citation:

Dahl TI and Nierenberg E (2021) Here's the TRIQ: The Tromsø Interest Questionnaire Based on the Four-Phase Model of Interest Development. *Front. Educ.* 6:716543. doi: 10.3389/feduc.2021.716543

The Tromsø Interest Questionnaire (TRIQ) is the first suite of self-report subscales designed for focused investigations on how interest is experienced in relation to Hidi and Renninger's four-phase model of interest development. In response to the plethora of varied interest measures that already exist in terms of theoretical grounding, form, and tested quality, the TRIQ subscales were designed with a consistent form to measure general interest, situation dependence, positive affect, competence level, competence aspirations, meaningfulness, and self-regulation answered in relation to some object of interest. Two studies testing the subscales' performance using different objects of interest (self-chosen "object-general," and prespecified "object-specific") provide evidence of the subscales' internal consistency, temporal reliability, and phase-distinguishing validity. Patterns across the two studies demonstrate that the TRIQ is a sufficiently reliable and valid domain-tailorable tool that is particularly effective at distinguishing phase 1 (triggered situational) from phase 4 (well-developed individual) interest. The findings raise interesting questions for further investigation about the distinction and distance between all interest phases, the push-pull factors that influence how interests evolve and additional subscales to add to the suite.

**Keywords:** interest, information literacy, four-phase model of interest development, reliability, validity, scale construction

## INTRODUCTION

We know it when we feel it, that feeling of interest in something. Sometimes the feeling hits us for the first time when we are introduced to that object or event in some attention-grabbing way, and sometimes that feeling is what drives us to pursue the interest more on our own. We know it when we feel it. But do we know it when we measure it?

## Interest

Interest catches and holds our attention (Hidi and Baird, 1986) and facilitates emotionally engaged interactions with objects of interest—critical activities for both initiating and sustaining learning over time (e.g., Harackiewicz et al., 2008). As a construct, interest has both emotional and cognitive aspects. Discrete emotions, or affective states, are perceptual and emotional processes that help prime us to focus on particular kinds of stimuli in involuntary physiological, preattentive ways that can vary in intensity and duration, and uniquely influence attention, behavior, and memory of the

stimuli involved (Dolan, 2002; Panksepp, 2003; Izard, 2011)—serving, in a sense, as “relevance detectors” (Scherer, 2005).

In the case of interest, the stimuli that catch our attention related to things we have less control over (novelty and uncertainty) set in motion curiosity-driven behaviors to explore them further (Silvia, 2006; Oatley et al., 2019). As a more conscious feeling, interest-related arousal—when cognitively appraised as something pleasant—becomes a driver for the deliberate pursuit of goals related to that particular object of interest (Scherer, 2005; Silvia, 2006), like cooking Thai food or solving math puzzles. Interest’s value lies in how, once triggered, it focuses our attention and orients us toward exploration and persistence in the face of obstacles (like having to find a rare ingredient or seeking harder math puzzles), uniquely fueling stamina and motivation more than other positive emotions such as enjoyment or happiness (Schiefele et al., 1992; Hidi et al., 2004; Thoman et al., 2011).

However, our relationship to each object or event can change over time—either evolving into a more stable interest (from interest to interests, as some describe, e.g., Berlyne, 1949; Silvia, 2006), or devolving into something situation dependent or no longer interesting (Hidi and Renninger, 2006).

Hidi and Renninger (2006) have been the most explicit in capturing this movement in their four-phase model of interest development. The model describes interest as an experience with four distinct developmental phases. These span experiences of triggered situational interest (phase 1), maintained situational interest (phase 2), emerging individual interest (phase 3), and well-developed individual interest (phase 4). The first two phases require more from the environment to initially trigger and maintain an interest (phase 1 being the most situation-dependent and fleeting), while interest in the last two phases is pursued increasingly more independently (phase 4 being the most independently pursued, and most stable in the face of obstacles).

To date, there has been no self-report tool that enables us to adequately test the experiences of Hidi and Renningers 2006 model in a unified manner. Renninger and Hidi (2011) have offered an overview of how interest has been operationalized and measured quantitatively and qualitatively by others, though without specific details of how those tools were developed and how they could ideally be used to test their model.

## Measuring Interest

If we were to rely on self-report measures that capture the underlying architecture of interest from Hidi and Renninger’s model, we must ground them in a clear definition of what interest is and put it in a form that is relatively quick and easy for people to describe in ways that are reliable and conceptually valid. Indeed, many have already developed self-report measures with this in mind. However, since our understanding of interest has changed over time, what has been measured in these self-report tools has varied, as has how the veracity of their value has been determined.

We began with a critical analysis of how other self-report measures have been developed in terms of 1) their theoretical grounding, 2) how they measure interest in domain tailorable ways, and 3) the evidence provided about existing measures’ reliability and validity. We found that though the concepts these

other tools touch on do overlap with aspects of the four-phase model of interest development, there is considerable variation among them, making none of them perfect matches for testing the full four-phase model—either alone or in combination with each other. Nevertheless, the way these other tools have been designed and tested is valuable to how we designed and tested ours to redress that gap. We therefore begin with an overview of other self-report interest measures and how that informed how we developed and tested the Tromsø Interest Questionnaire (TRIQ) subscales.

## Theoretical Bases

Naturally, over time, the bases for items and measures used have evolved along with how our understanding of interest has evolved (Renninger and Hidi, 2016). Preceding Hidi and Renninger’s four-phase model, self-report inventories were grounded in theories such as interest as an affective and dispositional state (Schiefele et al., 1988), the expectancy-value framework (Eccles et al., 1983; Eccles et al., 1993), interest as something that can be triggered, situational, and personal (e.g., Schraw et al., 1995; Ainley et al., 2002), interest as feeling and value (Krapp et al., 1988; Schiefele, 1999), conceptualization of interest as a multidimensional construct related to self-determination theory (Deci, 1992), interest as a part of the Cognitive-Motivational Process Model (Vollmeyer and Rheinberg, 2000), appraisal theory of interest (Silvia, 2006; Silvia, 2010), and self-concept theory (Marsh et al., 2005). Other measures have evolved with theories complementary to the four-phase model, such as work focused on the triggering, feeling, and value of interest (Linnenbrink-Garcia et al., 2010).

A small handful of measures have been designed more deliberately in harmony with the four-phase model. Examples of this work include that of Bathgate et al. (2014) who focused on the situatedness of interest, Ely et al. (2013) who focused on the stability of affect and interest over time, and Rotgans (2015) who focused on positive feelings, value, and the desire to reengage with an object of interest.

## Smorgasbord of Existing Measures

Existing measures of interest represent a varied terrain in both content and form that challenge the navigation of long lines of thought. To give a detailed sense of the variation, an overview of many of the commonly used measures is presented in **Supplementary Table S1**, with a brief overview here.

Of the simplest kinds of measures, participants can indicate their interests categorically with checklists, e.g., Bathgate et al. (2014), or dragging and dropping into categories (Ely et al., 2013). Single ratings of a selected set of objects of interest (such as topics or activities) have been used with 4-, 5-, 7-, and 10-point scales, e.g., Ainley et al. (2002), Alexander et al. (1995), Dawson (2000), and Häussler and Hoffmann (2002).

Scales with multiple ratings for each object of interest have also been developed with as few as 4 and as many as 24 items to rate on 4- to 7-point scales. They represent differential scale items (Silvia, 2010) and single factor scales such as the Individual Interest Questionnaire (IIQ, Rotgans, 2015), the Task Value and Competence Beliefs scales (Eccles et al., 1993), the Study

Interest Questionnaire (SIQ) and Cognitive Competence scale (Schiefele et al., 1988), the Situational Interest (SI) measures (Linnenbrink-Garcia et al., 2010), the Situational Interest Scale (SIS, Chen et al., 1999), parts of the Questionnaire on Current Motivation (QCM) (Vollmeyer and Rheinberg, 1998; Vollmeyer and Rheinberg, 2000), Affect and Experience scales (Ely et al., 2013), and others (e.g., Bathgate et al., 2014). Many of these scales are domain tailorable, though some were made specifically for particular domains, such as the Math Class-Specific Interest and Math Domain-Specific Interest measures (Marsh et al., 2005).

In sum, interest has been captured differently in terms of domain focus, the kinds of prompts and items used, the constellations of items used, and the scoring of the items used. Results from these measures are therefore difficult to compare, even for those that are conceptually related and domain tailorable. Additionally, because of the variety of questions and response alternatives, any attempt at combining these measures into a single scale would be cumbersome for respondents, hence the value of replacing these with a single comprehensive, domain-tailorable suite of measures.

## What has Been Used to Provide Evidence of Measure Reliability and Validity

Existing measures of interest vary in how much evidence has been provided regarding their reliability and validity. Though some do document the reliability of their interest measures, many do not. Among those that do, scale reliability has been tested and asserted using 1) internal consistency measures (Schiefele et al., 1988; Chen et al., 1999; Vollmeyer and Rheinberg, 2000; Häussler and Hoffmann, 2002; Marsh et al., 2005; Bathgate et al., 2014); 2) test-retest measures with lengths of delay as long as 2 months (Ely et al., 2013) and 3 months (Alexander et al., 1995); 3) confirmatory factor analysis to assert the construct reliability of a latent variable (Rotgans, 2015); 4) tests of multi-group invariance on a scale (e.g., Rotgans, 2015); and 5) Cohen's D to assert the reliability of a measure for distinguishing low- from high-scoring participants (Chen et al., 1999).

Fewer studies, though, provide evidence of the measure's validity. Those that do, have done so in several ways. Most commonly, evidence of construct validity has been tested with exploratory and confirmatory factor analyses (e.g., Eccles et al., 1993; Chen et al., 1999; Ainley et al., 2002; Linnenbrink-Garcia et al., 2010), yet other methods have also been used, including measures of face validity (e.g., Schraw et al., 1995), ecological validity (by, for example, relating interest to a classroom activity; Vollmeyer and Rheinberg, 2000), intraclass correlations (e.g., comparing relations between interest appraisals with appraisals of the objects of interest; Silvia, 2010), multi-group or multi-object invariance (e.g., Eccles et al., 1993; Chen et al., 1999; Linnenbrink-Garcia et al., 2010; Rotgans, 2015); predictive validity (Rotgans, 2015); convergent and divergent validity of interest measures correlated with other motivational measures (e.g., Schiefele et al., 1988; Bathgate et al., 2014); and structural equation modeling (e.g., Marsh et al., 2005).

If we want to test and develop a deeper understanding of the lived experience inherent to the multiple phases of interest Hidi

and Renninger have defined, we have to find ways to distinguish each phase from the others both conceptually and through people's reported interest experiences. That can be done by deliberately asserting and testing the unique psychological architecture underlying each phase. However, in our survey of the literature, we have yet to find a measure, or series of measures, that are sufficiently grounded in Hidi and Renninger's (2006) four-phase model of interest development to do that. If we can redress that, then we might also be able to unlock an even deeper understanding of how interest changes over time and how to better influence these changes. But first, it is time for a single, parsimonious tool with appropriate reliability and theoretically grounded construct validity to explore that (Kane, 2001). We also need a method for developing it so that the tool can be expanded in the future with additional subscales to help us test yet other aspects of that conceptual architecture.

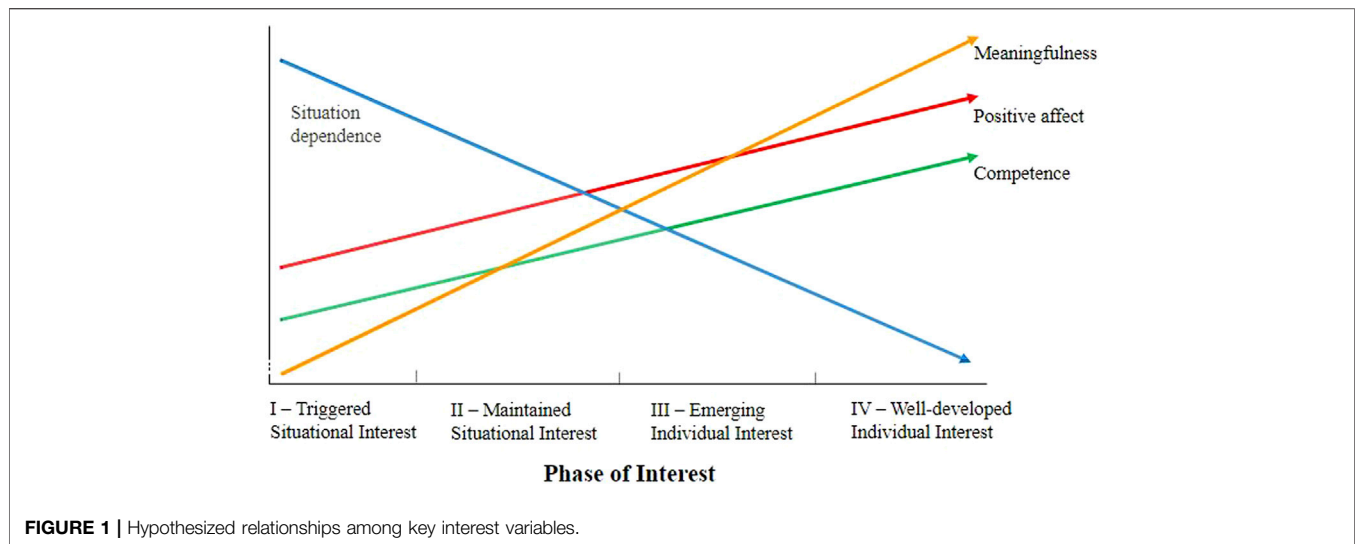
## The Posited Architecture of Each Phase

To set the stage for construct validity, our initial intent was to develop a tool that would capture general interest and four central elements of the four-phase model of interest development: situation dependence, positive affect, competence, and meaningfulness (Dahl, 2011; Dahl, 2014). If sufficiently robust, these measures should enable us to test the relationships we hypothesize among those elements at each phase of the model in ways that previous research yet has not done.

**Figure 1** summarizes the hypothesized relationships among four core variables from the four-phase model of interest development (Hidi and Renninger, 2006; Ekeland and Dahl, 2016; Dahl et al., 2019). These relationships are not posited to be absolute mean scores on a set scale, but rather relative relationships varying by degree between the core measured elements within and across each phase of interest. In other words, in line with Hidi and Renninger's model (2006), it is posited that the mean Situation Dependence score will be at its highest point in phase 1 and lowest in phase 4, steadily decreasing how much space for developing the interest is, by phase, situation-dependent. On the other hand, and also in line with the work of Linnenbrink-Garcia et al. (2010) and Schraw et al. (1995), we posit that the mean Positive Affect score in phase 1 will be lowest as a mean score in phase 1 (though higher than Competence and Meaningfulness), and steadily increase through to phase 4. Similarly, in line with Harackiewicz et al. (2008), competence will be lowest as a mean score in phase 1 and steadily increase with a similar slope to Positive Affect through to phase 4, though with a lower mean score at each phase than that of Positive Affect. Furthermore, in the spirit of findings from Bolkan and Griffin (2018), we posit that Meaningfulness mean scores will also be least and the lowest of all the four measured elements in phase 1, and steadily increase toward phase 4, exhibiting the steepest slope of all the increasing variables and highest mean score in phase 4.

## The TRIQ

Like the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1993), the TRIQ is being created as a modular suite of short subscales with a consistent form that can be expanded and tailored as needed in a conceptually and methodologically



streamlined way to better understand Hidi and Renninger's developmental model. What we ultimately hope to be able to answer with the TRIQ's subscales is how the quality of interest varies by interest phase.

### Item Bank

Prior to these studies, we created a large bank of items consisting of existing items from other measures, variations on existing items from other measures, and self-composed items. All had face validity in terms of the four key elements of the four-phase model of interest development. We later added items to this work to capture general interest and self-regulation, since reduced situation dependence implies increased demand for self-regulation. Based on principal component analyses of these items (unpublished manuscript), we selected sets of items for each scale that were found to be topically linked by underlying factors that represent situation dependence, positive affect, knowledge, and meaningfulness. From this preliminary work, we were motivated to divide "knowledge" into two variables—Competence Level (a respondent's assessment of what they currently know about, or are able to do with, the object of interest), and Competence Aspiration (a respondent's desire to know or be able to do more).

Finally, we wanted to be able to test our measure with students in Norway, and to create a measure that could also be used by the international research community. We therefore produced and tested the measure in both Norwegian and English. A pilot study yielded no significant differences between language forms (Nierenberg et al., 2021).

### Evidence of Reliability and Validity

We will provide evidence of reliability in the form of the internal consistency for each subscale using coefficient alpha; and temporal consistency using test-retest reliability with a 1-week delay and intraclass correlations. Evidence of phase-distinct construct validity will be tested with the subscales, by phase, with multivariate analyses of variance (MANOVA) and descriptive discriminant analysis (DDA). The findings will be

tested for multi-object invariance of all measures by comparing the results of one group who responded in terms of a self-chosen object of interest (the Object-General group) with the results from another group who responded in terms of a specified object of interest (the Object-Specific group).

### Study Design

Two studies were designed to test peoples' experiences of the two different target objects of interest: 1) self-chosen objects of interest ("Object-General," referred to as "X"), or 2) an object of interest that they were provided, namely interest in being or becoming an information literate person ("Object-Specific," referred to as "IL"). The first study was important for identifying phase similarities independent of the objects of interest upon which they were based. The second study was important for testing if those same patterns were consistent when all respondents were focused on one shared object of interest, in essence enabling us to compare the multiobject-invariance nature of the measures from the first study with the more person-invariant test from the second study. We chose IL as our specified object of interest since information literacy is a critical skill both in academic work (Feekey, 2013; Løkse et al., 2017) and in daily life, empowering people to be socially responsible consumers, users, and creators of information (Walton and Cleland, 2017).

In the Object-General questionnaire, participants read a short summary of each phase of the four phases of interest. They were then asked to identify one personal object of interest that fit each interest description. Participants were then randomly divided into even-sized groups by phase based on their birth months. Within their assigned phase, participants answered the scale questions about the self-identified X object of interest that they had listed for that phase. Though the objects of interest were unique to each participant within each phase, the quality of their interest in their object of interest was arguably similar.

In the Object-Specific questionnaire, we asked participants to indicate their phase of interest in a specific, predesignated IL

object of interest (in this case, being or becoming an information-literate person, with a clear definition provided). The participants in this round ended up in groups that were more varied in size, as people's interest in being or becoming an information-literate person naturally varies. Within each phase, however, the object of interest was the same for every participant.

## GENERAL METHODS

This section applies to both the Object-General (X) and Object-Specific (IL) investigations.

### Participants and Procedure

Data were collected between February 2019 and April 2020. Respondents answered the survey, twice, with either a time lapse of 1 week (test-retest) or one semester (pretest-posttest) between surveys. Whereas the test-retest group had no intervention, the pretest-posttest group received 2–4 hour of information literacy instruction between the surveys. The distinction between test-retest and pretest-posttest groups is therefore important for all analyses that involve Time 2 (T2) data.

#### Object-General (X) Participants

Data for the Object-General (X) measure were collected in both 2019 (X19) and 2020 (X20). In the combined T1 X-cohorts there were 335 participants (115 males, 215 females, 5 other). The 2019 participants ( $n = 86$ ) received the questionnaire once, while the X20 test-retest group ( $n = 247$ ) received the same questionnaire twice, with a 1-week time interval. Of those, 118 (35 males, 79 females, 1 other, 3 missing) also answered at T2. In both the rounds, ages ranged from 17 to 84 years (mode = 18–24, median = 35–44).

#### Object-Specific (IL) Participants

Data for the Object-Specific (IL) measure were collected in 2020. Two groups, with varying time intervals between T1 and T2, answered the survey in which participants' interest in being or becoming information literate (IL) was the specified object of interest. At T1, the two IL groups involved 364 participants all together (129 males, 229 females, 4 other, 2 missing). The age range of this combined group is 18–85 years, with a mean age of 23.1 years, and median and mode age spans of 18–24 years. At T2, the test-retest group analyzed in this study involved 69 participants (29 male, 40 female). The age range of this group is 18–84 years, with a mean age of 32.6 years, a median age span of 25–34 years, and a mode age span of 18–24 years.

The IL “test-retest group” ( $n = 253$  at T1) had only 1 week between surveys. This group consisted of both students and others, in Norway and the United States, who were recruited exclusively for this study. The IL “pretest-posttest group” ( $n = 111$  at T1) had nearly a whole semester, with 2–4 hour of IL-instruction between surveys, because they participated simultaneously in a longitudinal study tracking their IL-development over the semester. This group was comprised solely of first-year undergraduates in a wide range of

disciplines in Norway. Both the groups were combined in the T1 analysis, while only the test-retest group was used in the T2 analyses.

### Questionnaire Dissemination

The Object-General test-retest group's survey was distributed to a convenience sampling of Norwegian and English-speaking participants, through email and social media (e.g., mailing lists, FB, and Twitter). To secure heterogeneous age group representation (Etikan et al., 2016), we intentionally identified distribution sources that would enable us to reach a broad age range of participants in both Norway and the United States. Also, participants were encouraged to share the link to the survey with acquaintances. Furthermore, rewards were offered for those participants who answered also the second time. For this phase of our research, convenience sampling was feasible and sufficient, as the psychological architecture of interest development is regarded as a basic, neurologically based human experience (Hidi, 2006; Hidi and Renninger, 2006; Renninger and Hidi, 2011). Likewise, the questionnaire is in its early phase of development and open for additional scrutiny in subsequent work.

All test-retest participants were given the choice of answering in either Norwegian or English, while the pretest-posttest group answered the survey in Norwegian.

### Materials

An online interest survey, developed and distributed through Qualtrics, was used to collect the data (see **Supplementary Table S2**). Following questions about consent and demographics, the survey contained a “phase” section with four paragraphs, pretested for comprehensibility, which summarized the phases of interest derived from Hidi and Renninger's (2006) four-phase model of interest development. Participants assessed how well each interest phase description matched their target object of interest.

Following the phase section came the “subscale” section that contained a questionnaire comprised of the seven subscales that participants answered with respect to either their self-chosen (Object-General; X) or prespecified (Object-Specific; IL) object of interest. Once their general interest was assessed, the presentation of the remaining subscales was counterbalanced to avoid order effects. At the end of the survey, participants had an opportunity to write comments.

Below are example items from each subscale as presented in the Object-General questionnaire. “X” represented the interest that the respondent had self-identified for that phase. For the corresponding items in the Object-Specific questionnaire, X is replaced by “being or becoming an information literate person.” For example, for the X questionnaire, participants were asked “How interested are you in X” and for the IL questionnaire, participants were asked “How interested are you in being or becoming an information literate person.” Items in the General Interest and Positive Affect subscales utilized Likert scales ranging from 1 (not at all) to 6 (very much), while the other subscales used Likert scales ranging from 1 (not true at all) to 6 (very true).



- **General Interest:** Three items, modified from Renninger and Hidi (2011), including “How interested are you in X?”
- **Situation Dependence:** Three items, two self-composed and one adapted from Rotgans (2015), including “I am dependent upon others for maintaining my interest in X.”
- **Positive Affect:** Four matrix items, modified from Vittersø et al. (2005), including “How little or much do you experience these feelings (pleasure, happiness, interest, engagement) when you think about your interest in X?”
- **Competence Levels:** Three self-composed items, inspired by Rotgans (2015), including “I am satisfied with what I know about X.”
- **Competence Aspiration:** Three items, some self-composed and others adapted from Rotgans (2015) and Tracey (2002), including “I want to learn more about X.”
- **Meaningfulness:** Five items, some self-composed and some adapted from Rakoczy et al. (2005), Renninger and Hidi (2011), and Schiefele and Krapp (1996), including “Having an interest in X is very useful for me.”
- **Self-regulation:** Six self-composed items, including “I make time to develop my X-related knowledge and skills.”

## Analyses

In both the studies, how well the interest phase description matched their experience of their object of interest was assessed. Also, evidence of subscale reliability was determined through tests of *internal consistency* and *temporal consistency*.

The internal consistency of the items across all the phases was tested with Cronbach’s alpha for each subscale. Our preferred criterion for reliability was  $\alpha \geq 0.80$ , as this is generally considered good internal consistency (George and Mallery, 2003).

Temporal consistency, by phase and subscale, was tested in two ways: with test-retest analyses and intraclass correlations (ICC). Although interest is by definition somewhat fluid, and test-retest methods are typically used to determine a variable’s stability, a 1-week time interval was chosen as an interval long enough to limit memory of previous answers, yet short enough to limit the amount of interest change. Our reliability criterion for these tests was nonsignificant difference between Time 1 (T1) and Time 2 (T2) responses.

ICC analyses of T1 and T2 scores for each subscale by phase used to determine individual subscale scores indicated good agreement properties (Berchtold, 2016), by indicating whether within-individual scores are statistically similar enough to discriminate between individuals (Aldridge et al., 2017). We used ICC(A,N), a two-way random model with absolute agreement. Our three agreement criteria were that the ICC was positive, moderate to high ( $ICC \geq 0.50$ ), and significant ( $p < 0.05$ ) for each subscale by phase.

In terms of validity, analyses were done using all seven TRIQ subscales (General Interest, Situation Dependence, Positive Affect, Competence Level, Competence Aspiration, Meaningfulness, and Self-Regulation). Phase distinction was tested with MANOVA to determine if there were notable differences in the subscale scores by a person’s level (phase) of interest in their object of interest. Our reliability criterion was a significant main effect for subscale scores by phase, where all

subscale scores would increase from phase 1 to 4, apart from Situation Dependence, which would decrease. We had no a priori hypotheses about the size of the distinctions between phases.

Phase discriminant validity was tested with a descriptive discriminant function analysis with phase as the grouping variable, the subscales as the independent variables and prior probabilities set for all phases being equal (Huberty and Hussein, 2003; Warne, 2014; Barton et al., 2016). This enabled us to determine how well the interest phases are related to distinct interest experiences measured with the Time 1 subscales with the benefit of a reduced possibility of Type 1 error (Sherry, 2006), a method superior to multinomial logistic regression given our focus on the validity of the categories and the number of categories being tested (Al-Jazzar, 2012). Our validity criterion was one or more discriminant functions significantly correlated with the subscale scores in ways that distinguish people’s object of interest experience by phase. We also tested whether these distinctions predicted how accurately people’s experiences were categorized better than chance (chance being 25%).

Finally, the pattern of relationships was tested against the relationships posited in **Figure 1**.

All analyses were done using the statistical package IBM SPSS Statistics 26. In addition to what is reported in the results, the temporal reliability results and discriminant function classification tables are included in **Supplementary Table S3**.

## STUDY 1: OBJECT-GENERAL INTEREST

### Method

#### Materials

After the descriptions of the four phases of interest (see General methods), respondents were asked to identify, for each of the interest descriptions, one interest they had that matched the description (see “Object-General” in **Supplementary Table S3**). In addition to rating each self-provided object of interest by how well it matched with the relevant phase interest description (Match), they also rated how difficult it was to think of an appropriate example for each phase (Example-Finding Difficulty).

Based on their month of birth, respondents were then sent to the next part of the questionnaire, where they answered remaining questions in relation to the example they gave for one of the four interest phases, e.g., those born in the first 3 months described their phase 1 triggered situational interest example, those born in the second 3 months described their phase 2 maintained situational interest example, and so forth. What they chose is referred to as their self-chose object of interest “X.”

### Procedure

Participants who fully completed the questionnaire used a mean time of 16.9 min ( $SD = 13.4$ ) on the task. One week after completing the survey, the test-retest participants were sent a link, *via* e-mail, to the same questionnaire again.

In T2, participants were instructed to “write down which interest X you focused on in the first round. All your answers must be based on that interest again,” though we did not ask

which interest phase description that interest now matched. Unfortunately, due to a programming error in Qualtrics, the test-retest time lag for participants in phase four was 1 month instead of 1 week. These both have implications discussed in the results.

## Results

### What Kinds of Things did People Report Being Interested In?

The kinds of interests people answered their questions about varied from creative activity (29%, for example dance, photography, handiwork, art, food, music, writing), physical activity (20%, for example working out, cycling, skiing, scuba diving), intellectual activity (17%, for example computer programming, science, history), friluftsliv outdoor life (9%, for example fishing, hiking, hunting), and team sports (8%, for example soccer, handball). The remaining categories (5% or less each, included gardening, general maintenance, entertainment, social engagement, games, flying, and being social).

### Phase Description and Interest Match

To understand how well each interest phase description worked for stimulating object of interest examples and how well participant experience of those objects of interests fit with the interest phase description, mean scores were calculated for each.

For the Example-finding difficulty scores ( $n = 130$ ), the T1 means, standard deviations, and standard error of the mean were the following: phase 1,  $M = 3.63$ ,  $SD = 1.51$ ,  $SE = 0.13$ ; phase 2,  $M = 2.69$ ,  $SD = 1.41$ ,  $SE = 0.12$ ; phase 3,  $M = 2.29$ ,  $SD = 1.38$ ,  $SE = 0.12$ ; and phase 4,  $M = 1.79$ ,  $SD = 1.25$ ,  $SE = 0.11$ . Since higher scores indicated greater difficulty, these values suggest that it was progressively easier for participants to come up with examples for successively higher phases of interest ( $p < 0.01$  for all paired  $t$ -test comparisons).

For the Match scores ( $n = 130$ ), the T1 means, standard deviations, and standard error of the mean were the following: phase 1,  $M = 4.42$ ,  $SD = 1.18$ ,  $SE = 0.10$ ; phase 2,  $M = 4.70$ ,  $SD = 1.03$ ,  $SE = 0.09$ ; phase 3,  $M = 4.99$ ,  $SD = 1.07$ ,  $SE = 0.10$ ; phase 4,  $M = 5.25$ ,  $SD = 1.15$ ,  $SE = 0.10$ . A MANOVA of match by phase did not reveal any significant differences by phase,  $F(12, 375, 000) = 0.280.83$ ,  $ns$ , Pillai's trace = 0.027,  $\eta^2 = 0.165$ . This suggests that participants were likely thoughtful about choosing their examples to match the descriptions as best they could, providing a reasonable foundation for the next phase-based analyses.

### Evidence of Reliability

**Internal consistency: Cronbach's alpha.** The internal consistency of subscale items, measured with Cronbach's alpha ( $\alpha$ ), was calculated for each subscale using mean scores from T1. All but one subscale met our  $\alpha \geq 0.80$  criterion: General Interest ( $\alpha = 0.87$ ), Situation Dependence ( $\alpha = 0.84$ ), Positive Affect ( $\alpha = 0.81$ ), Competence Levels ( $\alpha = 0.76$ ), Competence Aspiration ( $\alpha = 0.91$ ), Meaningfulness ( $\alpha = 0.89$ ), and Self-regulation ( $\alpha = 0.90$ ).

**Paired sample t-tests by phase.** This test compared each subscale's test and retest scores for participants who completed both T1 and T2 questionnaires (see **Supplementary Table S3**).

Our reliability criterion was fulfilled for phases 1-3, exhibiting nonsignificant difference over time. Phase 4, however, showed significantly reduced means at T2 for four of the seven subscales ( $p < 0.05$  for General Interest, Competence Aspiration, Meaningfulness, and Self-regulation), and therefore only partially fulfilled this reliability criterion.

**Intraclass correlation by phase.** This test correlated T1 and T2 results for each subscale to determine the degree to which participants in the same phase score similarly both times. Our three reliability criteria were fulfilled for phases 1-3 for all subscales; the ICC was positive, moderate to high ( $ICC \geq 0.50$ ), and significant ( $p < 0.05$ ) (see **Supplementary Table S3**). However, only two subscales had significant correlations for phase 4 (Competence Level, 0.48\*\*, and Self-regulation, 0.38\*), yet neither of these ICC's were  $\geq 0.50$ . Our reliability criteria were therefore only partially fulfilled for phase 4.

### Evidence of Validity

**Phase distinction MANOVA.** A MANOVA run with T1 subscale scores by phase showed a main effect for phase,  $F(21, 717, 00) = 6.10$ ,  $p < 0.001$ , Pillai's Trace = 0.455,  $\eta^2 = 0.15$  (see **Table 1**). All but Situation Dependence showed significant increases by phase (all  $p < 0.001$ ). Within the subscales with distinct changes, there were nevertheless some nonsignificant differences between phases. Least significant difference post hoc tests showed nonsignificant differences between phases for the following subscales: General Interest, phases 2 and 3; Positive Affect, phases 3 and 4; Competence Level, phases 1 and 2 and phases 2 and 3; Competence Aspiration, phases 2, 3, and 4; Meaningfulness, phases 2, 3, and 4; Self-regulation, phases 2 and 3 and phases 3 and 4. Overall, then, for interest in X, the subscale scores provided the clearest distinction between Phases 1 and 4, and the least distinction between neighboring Phases 2 and 3, and 3 and 4.

**Discriminant function analysis.** For this analysis, the seven subscales were used to determine if participants' experience of interest qualitatively differed in meaningful ways by phase. Indeed, results show that the phases are associated with significantly distinct experiences.

Three discriminant functions were calculated. Function 1 had an effect size of 33.99% and Function 2 had an effect size of 8.41% (see **Table 2**). The first two discriminant functions accounted for 81% and 14% of the between-phase variability. Since the test of Function 3 was not significant, it is not included in the remainder of these analyses.

Standardized discriminant function coefficients and structure coefficients were examined to determine how the subscale variables contributed to the differences between phases (see **Table 3**). As the squared, pooled, within-group correlations between the subscales and canonical discriminant functions ( $r_s^2$ ) indicate, for Function 1, all of the variables aside from Situation Dependence significantly contributed to the group differences ( $p < 0.05$ ), though Competence Level much less so. For Function 2, only Competence Level was substantially responsible for phase differences.

As seen in **Figure 2**, the group centroids for each phase indicate that Function 1 maximally distinguishes phase 1 from

**TABLE 1** | Means, standard deviations, and multiple analysis of variance between-group effects for TRIQ object-general (X) subscales by interest phase.

TRIQ subscale	Phase 1 (n = 28)		Phase 2 (n = 66)		Phase 3 (n = 69)		Phase 4 (n = 84)		F(3,83)	$\eta^2$
	M	SD	M	SD	M	SD	M	SD		
General Interest	3.37	1.02	4.80	0.97	4.95	0.83	5.32	0.94	31.48	0.28
Situation Dependence	3.18	1.56	2.88	1.72	2.39	1.26	2.34	1.48	3.43	0.04
Positive Affect	3.66	1.18	4.51	0.81	4.87	0.91	5.02	0.89	17.30	0.18
Competence Level	3.29	1.37	3.73	1.08	3.81	1.07	4.24	0.98	6.44	0.07
Competence Aspiration	3.29	1.32	4.94	1.07	5.11	0.75	5.11	1.20	22.99	0.22
Meaningfulness	3.06	1.05	4.86	0.90	4.98	0.90	5.07	0.98	34.87	0.30
Self-Regulation	2.20	0.86	4.04	1.07	4.33	1.18	4.40	1.18	29.75	0.27

**TABLE 2** | Wilks' Lambda and canonical correlation for the four phases of interest in X.

Function	Wilks' Lambda	$\chi^2$	Df	P	Canonical correlation $R_c$	Effect size $R_c^2$
1-3	0.586	128,471	21	0.001	0.583	33.99%
2-3	0.997	28,744	12	0.004	0.290	8.41%
3	0.969	7,550	5	0.183	0.176	3.01%

**TABLE 3** | Standardized discriminant function and structure coefficients for the four phases of interest in X.

Subscale	Function 1			Function 2		
	Coefficient	$r_s$	$r_s^2$	Coefficient	$r_s$	$r_s^2$
General Interest	0.313	0.859*	73.79%	1.049	0.317	5.39%
Situation Dependence	0.065	-0.231	5.33%	-0.269	-0.240	5.76%
Positive Affect	-0.046	0.617*	43.19%	0.268	0.296	8.76%
Competence Level	0.156	0.329	10.82%	0.528	0.508*	25.81%
Competence Aspiration	0.116	0.736*	54.17%	-0.243	-0.238	5.66%
Meaningfulness	0.471	0.911*	82.99%	-0.667	-0.207	4.28%
Self-Regulation	0.248	0.842*	70.90%	-0.613	-0.097	1.40%

Note: The largest absolute correlation between each subscale and each discriminant function is indicated with an asterisk.

4, and Function 2 maximally distinguishes phase 2 from 4. The Function 1 and 2 centroids are for phase 1, -1.938 and 0.184; phase 2, 0.042 and -0.327; phase 3, 0.203 and -0.210; and phase 4, 0.447 and 0.368. This suggests that phase 1 is most distinct in relation to Function 1, particularly and in this order of effect (see Table 3), with less experience of Meaningfulness, General Interest, Self-regulation, Competence Aspiration, and Positive Affect. In terms of Function 2, phase 4 is slightly more distinct with a higher level experience of competence.

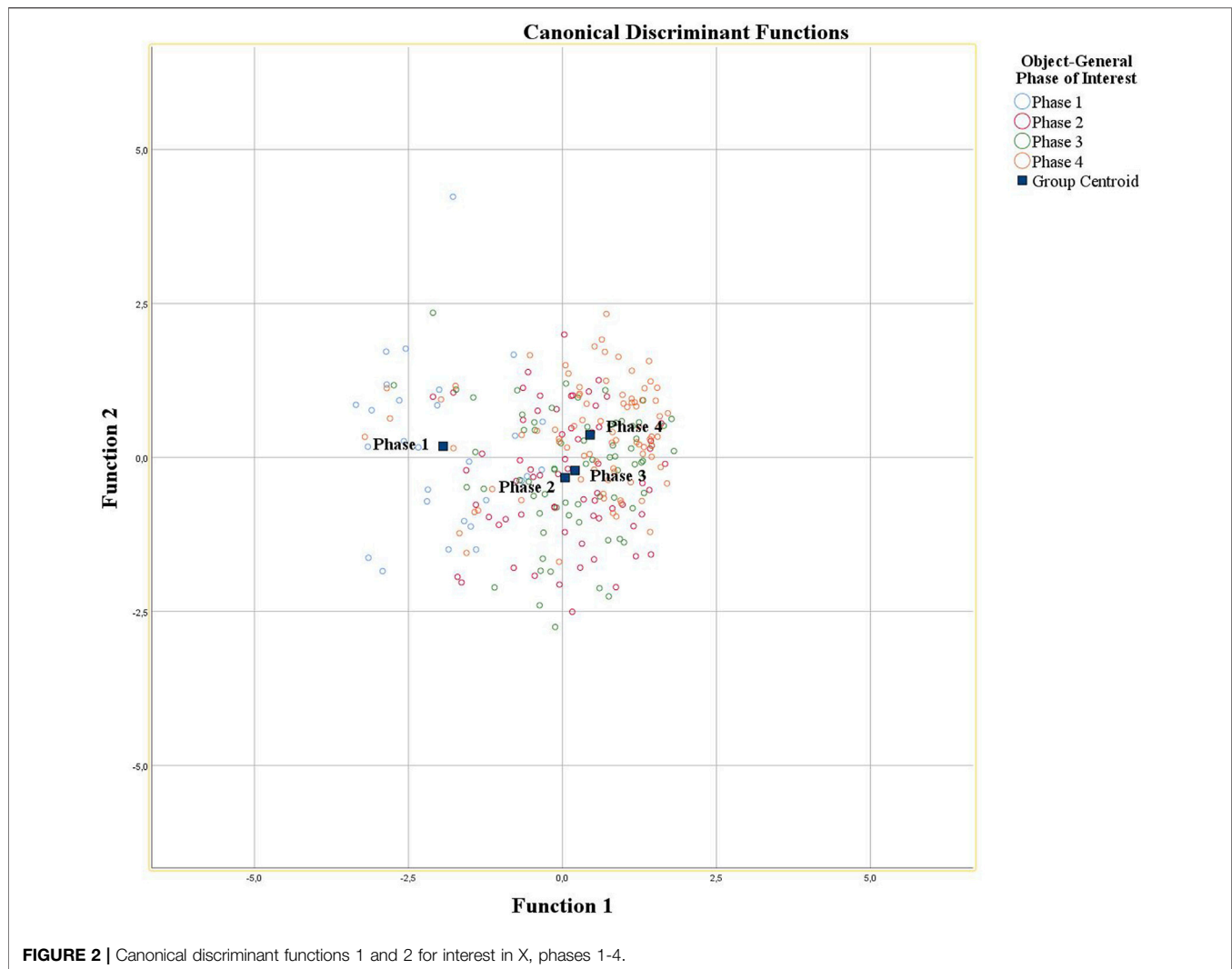
Based on the subscale hit rate for each self-reported interest phase, 51.5% of the original grouped cases were correctly classified in T1 by Function 1. As seen in Supplementary Table S3, the classifications were most distinct for phases 1 (75%) and 4 (58%), and less so for phases 2 (33%) and 3 (29%).

## Object-General Discussion

An advantage of this study is that the phase groups were designed to be relatively equal in size, though with varied objects of interest. As hoped, the match between how interested participants were in their self-chosen objects of interest and the interest phase descriptions was comparable across all phases. Furthermore, Cronbach's alpha indicated that all subscales showed strong

internal consistency. Six subscales met our criterion of  $\alpha \geq 0.80$ , and the seventh was only slightly below.

For the test-retest procedures, we recognize that given the number of *t*-test comparisons and varying sample sizes, this could increase likelihood of Type II errors. We therefore assess the overall pattern of results and give due attention to variation. First, the notable consistency with our predictions was promising. The additional ICC results that were also reasonably consistent with our predictions for phases 1-3 are a positive, corroborative point. However, variations in the data indicate more shifting than presumed, particularly for phase 4 interests. Most likely, this is because of the error in the data collection which gave T1 phase 4 respondents a greater lag time before responding to the T2 retest. This observed instability over time is as likely a result of the fluid nature of interest as the properties of the questionnaire itself—a point corroborated by the low intraclass correlations, though this remains to be verified with T1-T2 *t*-test comparisons. Since we did not ask which phase people were in with their object of interest at T2, we interpret the results for all phases with some caution. Note: to correct for the T2 phase presumption in subsequent research, we specifically asked for which phase participants in the object-specific study would classify their object of interest in at both T1 and T2.



The MANOVA and descriptive discriminant function analyses based on T1 data show significant differences in subscale scores by phase. Meaningfulness, General Interest and Self-regulation are most distinct across phases, followed by Competence Aspiration and Positive Affect. Given how distinctly General Interest contributes to this finding, we suggest using the General Interest subscale together with the other TRIQ subscales in future interest development research to account for other general aspects of interest not yet captured by the other variables. Given how much additional variance is accounted for by the other subscales, though, we tentatively suggest not using General Interest by itself when interest is the primary variable of study.

Based on the two discriminant functions, all of the variables, aside from Situation Dependence, were responsible for group differences captured by Function 1, though Competence Level contributed substantially to more phase differences accounted for by Function 2. Furthermore, correct phase classification was above chance in all cases, particularly for phases 1 and 4. Altogether, this indicates

that the combined subscales offer a moderately good, albeit uneven, indicator of how distinctly, and in which ways, each phase of interest is experienced.

That all the subscales combined worked best in distinguishing T1 phases 1 and 4 could be a result of the quality of the phase descriptions, the subscales themselves, or the fluid nature of interest development. In general, the greater ease people reported for finding later phase interest examples may indicate not only a deeper interest, but also greater metacognitive awareness of such interests. This may also have influenced the smaller number of phase 1 participants even though equally many people were randomly invited to each phase group. This is a matter to keep in mind in future collection and interpretation of self-report results.

In the end, the multi-object invariance related to the internal consistency of all subscales and the temporal stability of all subscales across phases 1, 2, and 3 is promising for TRIQ's research value. So, too, is the clear distinction of the T1 interest experience of phases 1 and 4. Nevertheless, the temporal stability of phase 4 remains to be properly tested.

## STUDY 2: OBJECT-SPECIFIC

### Method

#### Materials

In addition to the TRIQ focused on measuring students' interest in being or becoming information literate (see the General Methods section), the pretest-posttest group completed a 21-item multiple choice test measuring the students' knowledge of key aspects of information literacy, at both T1 and T2 (see Nierenberg et al., 2021). The presentation of these two questionnaires was evenly counterbalanced to limit possible order effects.

#### Procedure

The object-specific interest in being or becoming information literate (IL) test-retest group was recruited *via* social media and e-mail, using the same procedure as the object-general test-retest group. Participants used a mean of 14.5 min ( $SD = 10.8$ ) to fill out the T1 survey. Participants who completed the initial interest survey received a retest 1 week later *via* e-mail.

The pretest-posttest group was recruited from multiple disciplines at several higher education (HE) institutions in Norway where IL instruction was offered. The survey was distributed *via* the students' learning management system (LMS). A pretest (T1) was distributed at the beginning of the semester, before IL instruction, and an identical posttest (T2) was distributed 10–16 weeks later in the semester, after 2–4 hour of IL instruction. Rewards were provided to those who completed the posttest. Participants used a mean of 27.6 min ( $SD = 14.2$ ) to fill out the T1 survey with both the interest questionnaire and ILtest.

### Object-Specific Results

#### Phase Description and Interest Match

Mean scores were again calculated for Match. In this case, match indicated how well participants' experience of their interest in being or becoming information literate fit with each interest phase description.

For phases 1–4 ( $n = 354$ ), the T1 means and SDs for the phase description that matched their level of interest were as follows: phase 1,  $M = 4.37$  (1.06); phase 2,  $M = 4.79$  (0.83); phase 3,  $M = 4.84$  (1.04); phase 4,  $M = 5.37$  (0.78). These values show that the matches between participants' experiences of interest in being or becoming an IL person and the phase descriptions they identified as closest to that experience were moderate to strong for all phases (on the 1–6 scale)—though least for phase 1 and most for phase 4. A multivariate analysis of variance of match by phase was significant,  $F(12,1047) = 55.83$ ,  $p < 0.001$ , Pillai's trace = 1.17,  $\eta^2 = 0.39$ . Post hoc tests indicated that the Match score was significantly distinct, and consistently higher for phase 4 than for the other three phase descriptions. This provides a reasonable foundation for the next phase-based analyses.

#### Evidence of Reliability

**Internal consistency: Cronbach's alpha.** The internal consistency of subscales in Study 2's object-specific study was calculated using mean T1 subscale scores, with the same criterion as Study 1 of  $\alpha \geq 0.80$ . All subscales met this condition: General Interest ( $\alpha = 0.84$ ), Situation Dependence ( $\alpha = 0.80$ ), Positive Affect ( $\alpha = 0.87$ ),

**TABLE 4** | Phase of interest for participants at T1 (Test) and T2 (Retest).

T2 retest phase					
T1 test phase	Situational interest		Personal interest		
	Phase 1	Phase 2	Phase 3	Phase 4	Total
Phase 1	<b>0</b>	1	0	0	1
Phase 2	1	<b>8</b>	4	2	15
Phase 3	0	3	<b>16</b>	3	22
Phase 4	0	2	7	<b>20</b>	29
Total	1	14	27	25	67

Note: Bold values indicate the participants who reported being at the same phase of interest and T1 and the T2 test-retest phase.

Competence Level ( $\alpha = 0.89$ ), Competence Aspiration ( $\alpha = 0.91$ ), Meaningfulness ( $\alpha = 0.93$ ), and Self-regulation ( $\alpha = 0.88$ ).

**Temporal consistency.** To be considered reliable, each subscale should meet the four criteria detailed in the tests below. However, as motivated by the object-general study, the findings in **Table 4** indicate that during the week between the test and retest, some people's interest in being or becoming information literate shifted, mostly to a neighboring level, and mostly to a higher level of interest.

In light of the interests in phases 1 and 2 being more situational and in phases 3 and 4 as more personal, we note the following. For the shifters who classified their interest as being in phase 2 at T1, many reported a more personal interest in IL at the T2 retest. For those who classified their interest as being in phase 3 at T1, the interest was more stable, with an equally small number of shifters experiencing an increase to a more personal interest level and a decrease to a more situational interest level. The T1 phase 4 shifters experienced a decrease in interest, but still mostly within a level of personal interest.

The test-retest analyses capture the nature of an interest experienced within a particular phase. The subsequent analyses are therefore linked to the interests that were constant at T1 and T2 ( $n = 44$ ; bold in **Table 4**). Note that there are therefore no temporal results reported for phase 1—already a small group of one, and arguably the most tenuous of all the phases.

**Paired sample t-tests by phase.** This test compared subscale T1 and T2 data from the test-retest group. Our reliability criterion was fulfilled for phases 2–4, exhibiting consistent nonsignificant difference over time (see **Supplementary Table S3**). All test-retest scores for all scales were nonsignificant with only two exceptions: phase 3 and phase 4 Meaningfulness scores increased significantly over the week.

**Intraclass correlation by phase.** The same three ICC-criteria as in Study 1 were used to interpret correlations in Study 2, namely that ICC should be positive, moderate to high, and significant. All of the criteria were met for General Interest, Positive Affect, and Competence Aspiration (see **Supplementary Table S3**). In terms of the criteria met with the other scales, for Situation Dependence, phase 3 met all criteria, phase 4 met two, and phase 2 met one (though correlations for both phases 4 and 2 nevertheless moderate and therefore close). For Competence Level, phase 4 met all, phase 2 met two (though with a larger sample, the correlation would have been significant), and phase 3 met one. For Meaningfulness, phase 4 met two, phase 3 met one, and phase 2 did not meet any of the

**TABLE 5** | Means, standard deviations, and multiple analysis of variance between-group effects for TRIQ object-specific (IL) subscales by interest phase.

TRIQ subscale	Phase 1 (n = 22)		Phase 2 (n = 114)		Phase 3 (n = 143)		Phase 4 (n = 61)		F(3,336)	$\eta^2$
	M	SD	M	SD	M	SD	M	SD		
General Interest	2.97	0.98	4.06	0.83	4.62	0.82	5.24	0.76	52.62	0.32
Situation Dependence	3.11	1.17	2.97	1.21	2.54	1.00	2.33	1.14	6.58	0.06
Positive Affect	2.97	1.09	3.74	0.98	4.07	0.98	4.64	0.85	20.49	0.16
Competence Level	2.42	0.94	2.95	1.01	3.32	1.03	4.04	1.10	19.92	0.15
Competence Aspiration	3.35	1.97	4.30	1.20	4.69	1.00	4.91	0.94	14.36	0.11
Meaningfulness	3.48	1.11	4.36	0.85	4.67	0.88	5.24	0.78	26.30	0.19
Self-Regulation	2.60	0.85	3.19	0.86	3.78	0.89	4.70	0.81	53.01	0.32

Note: All F-tests are significant,  $p < 0.001$ .

**TABLE 6** | Wilks' Lambda and canonical correlation for the four phases of interest in IL.

Function	Wilks' Lambda	$\chi^2$	Df	p	Canonical correlation $R_c$	Effect size $R_c^2$
1–3	0.564	191,046	21	0.000	0.631	39.82%
2–3	0.937	21,587	12	0.042	0.218	4.75%
3	0.985	5,384	5	0.371	0.127	1.61%

criteria. Finally, for Self-regulation, phase 4 met all criteria, phase 3 met one (though the correlation was nevertheless moderate and therefore close), and phase 2 met none. Meaningfulness showed the greatest variation from T1 to T2 even though, overall, mean Meaningfulness scores increased from one phase to the next with relatively low variance.

### Evidence of Validity

**Phase distinction MANOVA.** A MANOVA run with T1 test-retest and pretest-posttest subscale scores by phase showed a main effect for phase,  $F(21,996) = 8.63$ ,  $p < 0.001$ , Pillai's Trace = 0.462,  $\eta^2 = 0.15$  (see **Table 5**). All subscales changed as predicted—all significantly increasing by phase, aside from Situation Dependence, which significantly decreased by phase. Within the subscales, least significant difference post hoc tests showed that almost all scores significantly differed from each other by phase in the predicted direction ( $p < 0.01$  for all and  $p < 0.001$  for most), with the following exceptions: 1) for Situation Dependence, neither phases 1 and 2 nor phases 3 and 4 were significantly distinct, and for phases 1 and 3,  $p < 0.05$ ; 2) for Competence Level, phase 1 and 2  $p < 0.05$ ; and finally, 3) for Competence Aspirations, phases 3 and 4 were not significantly distinct.

**Discriminant function analysis.** For this analysis, all the subscales were used to determine if the phase of interest that T1 participants in both the test-retest and pretest-posttest groups were experientially distinct.

Three discriminant functions were calculated, and showed a moderately high correlation with Function 1, with an effect size of 39.82%, and low canonical correlation with Function 2 (see **Table 6**). The first two discriminant functions accounted for 91 and 7% of the between-phase variability. The test of function 3 was not significant and therefore not considered in the remainder of these analyses.

Standardized discriminant function coefficients and structure coefficients were examined to determine how the subscale

variables contributed to the differences between phases (see **Table 7**). As the squared, pooled, within-group correlations between the subscales and canonical discriminant functions ( $r_s^2$ ) indicate, for Function 1, all of the variables aside from Situation Dependence significantly contributed to the group differences ( $p < 0.05$ ), though, in this case Competence Aspiration a bit less so. For Function 2, Competence Aspiration was primarily responsible for group differences. Note that, compared to the object-general analyses, the same variables were significantly associated with Functions 1 and 2. However, in the object-specific analyses, Competence Level and Aspiration traded which function they significantly correlated with.

Results showed that people were correctly classified by their subscale scores into their selected phase well beyond chance. Based on the subscale hit rates for each T1 self-reported interest phase, 49.4% of the original grouped cases were correctly classified by Function 1 in T1 (Wilks' Lambda = 0.56). The results of the discriminant function analyses are found in **Supplementary Table S3**.

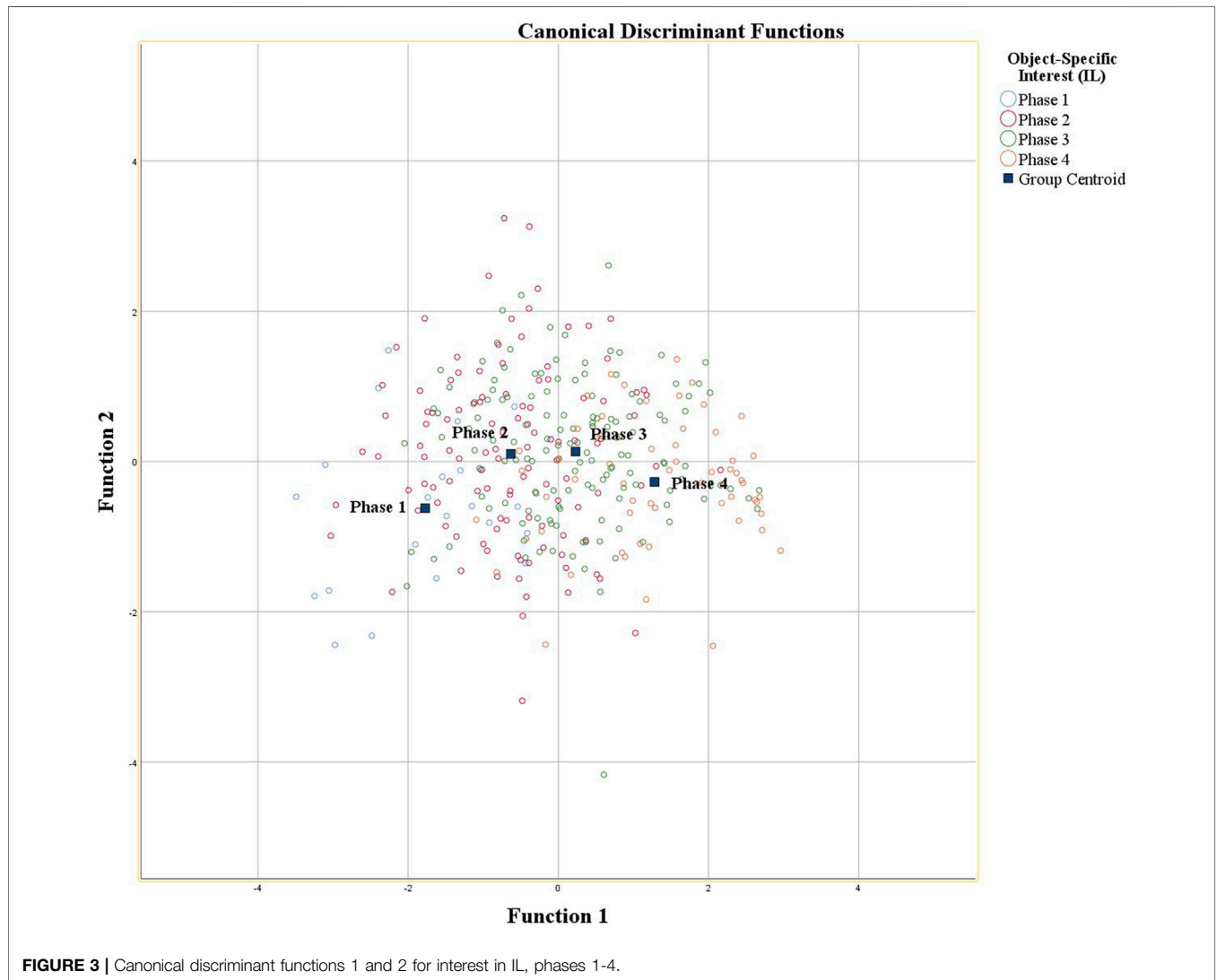
As seen in **Figure 3**, the group centroids for each phase indicate that Function 1 maximally distinguishes phase 1 from 4, and Function 2 maximally distinguishes phase 1 from phases 2 and 3. The Function 1 and 2 centroids are for phase 1,  $-1.773$  and  $-0.622$ ; phase 2,  $-0.632$  and  $0.100$ ; phase 3,  $0.230$  and  $0.132$ ; and phase 4,  $1.282$  and  $-0.273$ . This suggests that phase 1 is distinct in relation to Function 1, particularly and in this order of effect (see **Table 7**), in terms of lower General Interest and Self-regulation, and then Meaningfulness, Positive Affect, and Competence Level. In terms of Function 2, Phases 1 and 4 are distinguished by lower Competence Aspiration than phases 2 and 3.

### Object-Specific Discussion

While an advantage in this study is that the object of interest was the same for all participants, a disadvantage is that the number of

**TABLE 7 |** Standardized discriminant function and structure coefficients for the four phases of interest in IL.

Subscale	Function 1			Function 2		
	Coefficient	$r_s$	$r_s^2$	Coefficient	$r_s$	$r_s^2$
General Interest	0.694	0.834*	69.56%	0.790	0.426	18.15%
Situation Dependence	-0.146	-0.284	8.01%	0.056	0.027	0.10%
Positive Affect	-0.186	0.521*	27.14%	-0.081	0.111	1.23%
Competence Level	0.255	0.512*	26.21%	0.165	-0.228	5.20%
Competence Aspiration	0.044	0.411	16.89%	0.652	0.572*	32.72%
Meaningfulness	-0.155	0.585*	34.22%	-0.256	-0.220	4.84%
Self-Regulation	0.500	0.838*	70.22%	-0.959	-0.408	16.65%

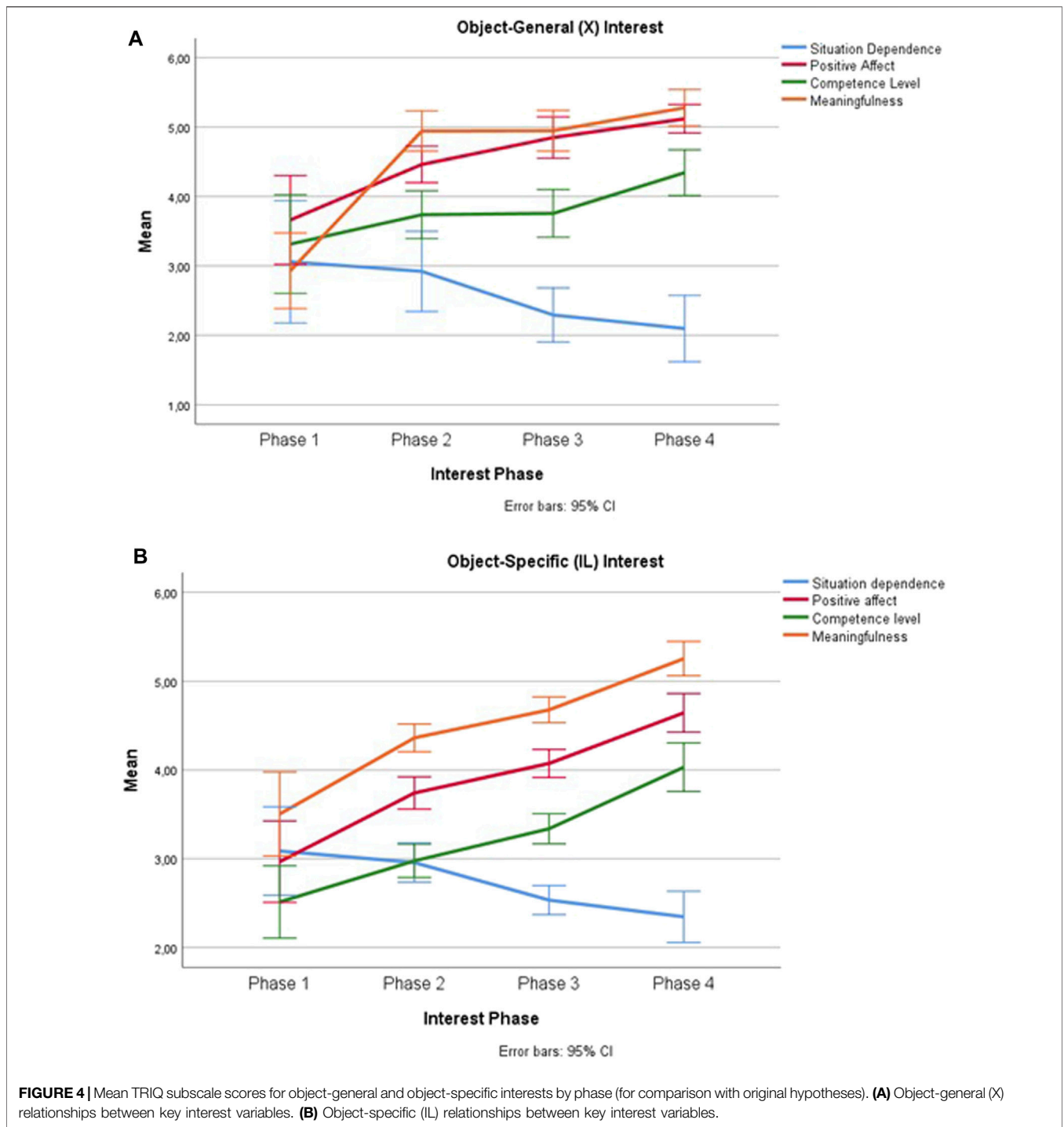


**FIGURE 3 |** Canonical discriminant functions 1 and 2 for interest in IL, phases 1-4.

participants in each phase varied, though larger participant groups helped compensate for that deficit.

The test-retest findings in Study 2 indicate, again, that interests have a fluid quality. We were able to catch people’s changes in interest by asking people at T1 and T2 which phase description best described their interest in being or

becoming information literate. The results indicated that even over the course of a week, interests can (d)evolve. **Table 4** indicates that regarding IL, the participants in this study tended to remain in or move into phases of individual interest more than situational interest. We therefore suggest that participants both indicate their general phase of interest



**FIGURE 4 |** Mean TRIQ subscale scores for object-general and object-specific interests by phase (for comparison with original hypotheses). **(A)** Object-general (X) relationships between key interest variables. **(B)** Object-specific (IL) relationships between key interest variables.

and fill out the TRIQ subscales when documenting the quality of their interest over time (which we had not done in the object-general study).

Though the ICC analyses showed the greatest consistency in participants’ General Interest, Positive Affect, and Competence Aspiration over time, where the subscales did show some variation, they varied least in phase 4, then in phase 3, and finally in phase 2. This pattern fits with what one

might predict about fluidity from the more established nature of individual interests (phases 3 and 4 being more stable, as expected, though unable to test in the object-general study because of the phase 4 snafu) over situational interests (phases 1 and 2 being more fluid) in the four-phase model. This also addresses the question we were unable to answer about Phase 4 in the object-general study because of the longer T2 delay.



Based on two discriminant functions, all of the variables aside from Situation Dependence were responsible for group differences captured by Function 1, though Competence Aspiration contributed substantially to more group differences accounted for by Function 2. Actual interest experiences and interest description matches were strongest for phases 1 and 4, though in all the cases, the discriminant function analysis demonstrated phase prediction above chance. Again, this indicates that the combined subscales offer a moderately good, albeit uneven, indicator of how distinctly, and in which ways, each phase of interest is experienced. Furthermore, the combined strength of all the subscales reinforces the value of including the General Interest subscale to capture aspects of interest otherwise not measured by the other TRIQ subscales, and vice versa when interest phase or development is the primary object of the study.

## OBJECT-GENERAL AND OBJECT-SPECIFIC VARIABLE RELATIONSHIPS

When we compare the relationships among the core variables in the original hypotheses illustrated in **Figure 1** with both the object-general and object-specific results represented in **Figures 4A,B**, the patterns of means (and confidence intervals, CI) are notably similar, and corroborate the findings from the MANOVA and descriptive discriminant analyses.

In absolute terms, the original variables we set out to study, in both the object-general and object-specific analyses, indicated, as predicted, that Positive Affect, Competence Level, and Meaningfulness all increase significantly and in distinct ways by phase. However, Situation Dependence varied less by phase than predicted. Also, the Meaningfulness score deviated from our phase 1 prediction in the object-specific study by being higher than Positive Affect and Competence Level. This is perhaps an artifact of that particular object of interest (IL) in as probed among participants from academic settings. The degree to which a triggered interest can be immediately meaningful to people may be more situation-dependent than we had hypothesized. This warrants further research.

Meanwhile, as already reported, the variables of General Interest, Competence Aspiration, and Self-regulation that were added after the original hypotheses were posited also increased significantly by phase, as one would also expect.

That these patterns were quite similar in both studies, and that most of the subscales correlated significantly with the first discriminant function, suggests that the subscales capture important aspects of interest in a coherent and arguably domain-tailorable way, providing compelling evidence for construct validity (Kane, 2001). Accordingly, we suggest that TRIQ can be used in future research to study general and specified interests, though the interaction between specified interests and context may offer additional, important information about how an interest devolves.

## OVERALL DISCUSSION

Have we designed a tool that can help us study the four different phases of interest described by Hidi and Renninger (2006) in reliable and valid ways?

We have designed subscales that are uniform in design and have documented their internal consistency and temporal reliability. In terms of inferences we can make about the subscales' construct validity relevance, as originally posited, evidence indicates qualitatively different experiences in predictable ways in terms of each phase's absolute and relative subscale score means and slopes.

What this evidence does not indicate, however, is equal distance among the variables by phase, suggesting that the difference between them is less defined, at least by the key variables we focused on, than asserted in the four-phase model. However, the General Interest scale captures differences between the phases that have yet to be distinguished, so understanding more about what the General Interest measure captures and explicitly distinguishing any additional underlying factors with their own subscales may help define the phases more distinctly. For example, we suggest the TRIQ suite of subscales be supplemented with a subscale to measure the desire to reengage (Rotgans, 2015) and perhaps a general self-efficacy measure for problem-solving—perhaps an important factor for realizing the desire to pursue an interest further on one's own (e.g., Chen et al., 2001).

## Limitations and Suggestions for Future Research

We found that the TRIQ subscales did not reveal equally distinct interest experiences for all interest phases. The blurrier distinction between phases 2 and 3 and phases 3 and 4, may suggest that either our tools are not sharp enough to discern these distinctions, or the psychological distance between them may actually be less than the psychological distance between phase 1 and all the rest. The next step would be to determine if 1) that is an artifact of people describing interests from the past or present, or 2) if there are additional variables that ought to be tested and included in the TRIQ to better distinguish people's phase 2-4 interest experiences.

Also, our work was originally motivated by the nature of general interest and the four phases of interest as they related to situation dependence, positive affect, knowledge, and meaningfulness. However, in preparing the items for our measures, we ended up replacing knowledge with two more distinct knowledge measures (Competence Level and Competence Aspiration) and added a Self-regulation measure. These additions accounted for additional variance and are worthy of closer scrutiny.

Additionally, recruiting and retaining sufficient numbers of participants for repeated measures work poses unique challenges, hence our need to recruit multiple samples of participants over time. This is worth keeping in mind in the design of future subscale contributions to the TRIQ suite of measures.

Finally, in the interest of better understanding how interest (d) evolves, it would be useful to employ the TRIQ to test what moves people from one phase to the next and if there are border-distinct push or pull factors between phases. For example, what is the impact of relevance (where meaningfulness and competence

intersect) or resonance (where meaningfulness and positive affect intersect) on one's draw to a higher or lower phase? How might other questions and methods supplement TRIQ findings to help us understand those relationships even better? For example, by evaluating the probabilities of being associated with one of two neighboring categories using multinomial logistic regression (DeRose, 1991; Al-Jazzar, 2012) or by elucidating the experience of particular phases of interest through qualitative interviews (Renninger and Hidi, 2011)?

## CONCLUSION

The object-general study allowed us to test how stable the TRIQ scales were, by phase, across various domains, while the object-specific study allowed us to retest the veracity of the object-general findings in an object-specific way. The similarities between the two studies' findings offer compelling evidence that the TRIQ suite of subscales are reliable, theoretically valid, and can therefore be useful for studying varying and phase-distinct experiences of interest commensurate with Hidi and Renninger's (2006) four-phase model of interest development. We encourage the development of more subscales to add to the TRIQ suite, adding variable measures that use the same basic form, and tested similarly for reliability and validity.

As for interest, we still know it when we feel it. Can we now get to know it even better with this new way of measuring it? It is a TRIQ question.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

## REFERENCES

- Ainley, M., Hidi, S., and Berndorff, D. (2002). Interest, Learning, and the Psychological Processes That Mediate Their Relationship. *J. Educ. Psychol.* 94 (3), 545–561. doi:10.1037/0022-0663.94.3.545
- Al-Jazzar, M. (2012). *A Comparative Study between Linear Discriminant Analysis and Multinomial Logistic Regression in Classification and Predictive Modeling*. Gaza: Master's thesis, Al Azhar University Gaza.
- Aldridge, V. K., Dovey, T. M., and Wade, A. (2017). Assessing Test-Retest Reliability of Psychological Measures. *Eur. Psychol.* 22, 207–218. doi:10.1027/1016-9040/a000298
- Alexander, P. A., Jetton, T. L., and Kulikowich, J. M. (1995). Interrelationship of Knowledge, Interest, and Recall: Assessing a Model of Domain Learning. *J. Educ. Psychol.* 87 (4), 559–575. doi:10.1037/0022-0663.87.4.559

institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

TID was primarily responsible for the conception and the design of study 1, while TID and EN contributed equally to the conception and design of study 2. TID and EN organized the databases for both studies. EN primarily organized and performed the descriptive analyses while TID organized and performed the inferential statistical analyses. TID and EN both wrote sections of the manuscript, and TID completed the first full draft. Both authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work was initially funded by The Research Council of Norway for the project Service Innovation and Tourist Experiences in the High North 195306/I40 (2009–2017). The publication of this article is funded by UiT The Arctic University of Norway.

## ACKNOWLEDGMENTS

The authors would like to thank the Norwegian Research Council for funding the Northern InSights project 195306/i40 that helped initiate this work. They would also like to thank UiT students Ingvild Lappegård, Karine Thorheim Nilsen, Torstein de Besche, Tilde Ingebrigtsen and Jesper Kraugerud, along with early contributors Kjørsti Thorsteinsen and Tony Pedersen. Their involvement in the project along the way was critical to its healthy start and final completion.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2021.716543/full#supplementary-material>

- Barton, M., Yeatts, P. E., Henson, R. K., and Martin, S. B. (2016). Moving Beyond Univariate Post-hoc Testing in Exercise Science: A Primer on Descriptive Discriminate Analysis. *Res. Q. Exerc. Sport.* 87 (4), 365–375. doi:10.1080/02701367.2016.1213352
- Bathgate, M. E., Schunn, C. D., and Correnti, R. (2014). Children's Motivation Toward Science Across Contexts, Manner of Interaction, and Topic. *Sci. Ed.* 98 (2), 189–215. doi:10.1002/sce.21095
- Berchtold, A. (2016). Test-retest: Agreement or Reliability? *Methodological Innov.* 9, 205979911667287. doi:10.1177/2059799116672875
- Berlyne, D. E. (1949). Interest as a Psychological Concept. *Br. J. Psychol. Gen. Sect.* 39 (4), 184–195. doi:10.1111/j.2044-8295.1949.tb00219.x
- Bolkan, S., and Griffin, D. J. (2018). Catch and Hold: Instructional Interventions and Their Differential Impact on Student Interest, Attention, and Autonomous Motivation. *Commun. Education.* 67 (3), 269–286. doi:10.1080/03634523.2018.1465193

- Chen, A., Darst, P. W., and Pangrazi, R. P. (1999). What Constitutes Situational Interest? Validating a Construct in Physical Education. *Meas. Phys. Education Exerc. Sci.* 3 (3), 157–XXX. doi:10.1207/s15327841mpee0303\_3
- Chen, G., Gully, S. M., and Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Res. Methods.* 4 (1), 62–83. doi:10.1177/109442810141004
- Dahl, T. I., Lappegård, I., Thorheim, K., and Nierenberg, E. (2019). “Developing a Measure of Interest: Where to Next From This fork in the Road?,” in 20th Norwegian Social Psychology and Community Psychology Conference Tromsø, Norway.
- Dahl, T. I. (2011). “Moving People: A Conceptual Framework for Capturing Important Learning From High North Experiences,” in *Opplevelser I Nord/ Northern InSights*. Svolvær, Norway.
- Dahl, T. I. (2014). “Moving People: A Conceptual Framework for Understanding How Visitor Experiences Can Be Enhanced by Mindful Attention to Interest,” in *Creating Experience Value in Tourism*. Editors N.K. Prebensen, J.S. Chen, and M. Uysal (Boston, MA: CABI), 79–94. doi:10.1079/9781780643489.0079
- Dawson, C. (2000). Upper Primary Boys’ and Girls’ Interests in Science: Have They Changed Since 1980? *Int. J. Sci. Education.* 22 (6), 557–570. doi:10.1080/095006900289660
- Deci, E. L. (1992). “The Relation of Interest to the Motivation of Behavior: A Self-Determination Theory Perspective,” in *The Role of Interest in Learning and Development*. Editors K.A. Renninger, S. Hidi, and A. Krapp (Hillsdale, NJ, USA: Lawrence Erlbaum), 43–69.
- DeRose, D. (1991). *Comparing Classification Models From Multinomial Logistic Regression and Multiple Discriminant Analysis*. Ann Arbor, MI: University Microfilms International.
- Dolan, R. J. (2002). Emotion, Cognition, and Behavior. *Science.* 298 (5596), 1191–1194. doi:10.1126/science.1076358
- Eccles, J., Wigfield, A., Harold, R. D., and Blumenfeld, P. (1993). Age and Gender Differences in Children’s Self- and Task Perceptions During Elementary School. *Child. Dev.* 64 (3), 830–847. doi:10.2307/113122110.1111/j.1467-8624.1993.tb02946.x
- Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., et al. (1983). “Expectancies, Values, and Academic Behaviors,” in *Achievement and Achievement Motives*. Editor J. T. Spence (San Francisco: Freeman), 75–146.
- Ekeland, C. B., and Dahl, T. I. (2016). Hunting the Light in the High Arctic: Interest Development Among English Tourists Aboard the Coastal Steamer Hurtigruten. *Tourism Cult. Commun.* 16 (1–2), 33–58. doi:10.3727/109830416X14655571061719
- Ely, R., Ainley, M., and Pearce, J. (2013). More Than Enjoyment: Identifying the Positive Affect Component of Interest That Supports Student Engagement and Achievement. *Middle Grades Res. J.* 8 (1), 13–32.
- Etikan, I., Musa, S. A., and Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *Ajtas.* 5 (1), 1–4. doi:10.11648/j.ajtas.20160501.11
- Feekery, A. J. (2013). *Conversation and Change: Integrating Information Literacy to Support Learning in the New Zealand Tertiary Context: A Thesis Presented for Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy at Massey University*. Manawatu: Massey University.
- George, D., and Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update*. Boston, MA, USA: Allyn & Bacon.
- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., and Tauer, J. M. (2008). The Role of Achievement Goals in the Development of Interest: Reciprocal Relations Between Achievement Goals, Interest, and Performance. *J. Educ. Psychol.* 100 (1), 105–122. doi:10.1037/0022-0663.100.1.105
- Häussler, P., and Hoffmann, L. (2002). An Intervention Study to Enhance Girls’ Interest, Self-Concept, and Achievement in Physics Classes. *J. Res. Sci. Teach.* 39 (9), 870–888. doi:10.1002/tea.10048
- Hidi, S., and Baird, W. (1986). Interestingness-A Neglected Variable in Discourse Processing. *Cogn. Sci.* 10 (2), 179–194. doi:10.1016/S0364-0213(86)80003-9
- Hidi, S. (2006). Interest: A Unique Motivational Variable. *Educ. Res. Rev.* 1 (2), 69–82. doi:10.1016/j.edurev.2006.09.001
- Hidi, S., Renninger, K. A., and Krapp, A. (2004). “Interest, a Motivational Variable that Combines Affective and Cognitive Functioning,” in *He Educational Psychology Series. Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development*. Editors D. Y. Dai and R. J. Sternberg (Lawrence Erlbaum Associates Publishers), 89–115. doi:10.4324/9781410610515-11
- Hidi, S., and Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educ. Psychol.* 41 (2), 111–127. doi:10.1207/s15326985ep4102\_4
- Huberty, C. J., and Hussein, M. H. (2003). Some Problems in Reporting Use of Discriminant Analyses. *J. Exp. Education.* 71 (2), 177–192. doi:10.1080/00220970309602062
- Izard, C. E. (2011). Forms and Functions of Emotions: Matters of Emotion-Cognition Interactions. *Emot. Rev.* 3 (4), 371–378. doi:10.1177/1754073911410737
- Kane, M. T. (2001). Current Concerns in Validity Theory. *J. Educ. Meas.* 38 (4), 319–342. doi:10.1111/j.1745-3984.2001.tb01130.x
- Krapp, A., Schiefele, U., and Winteler, A. (1988). Studieninteresse und fachbezogene Wissensstruktur. *Psychol. Erziehung Unterricht.* 35, 106–118.
- Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., et al. (2010). Measuring Situational Interest in Academic Domains. *Educ. Psychol. Meas.* 70 (4), 647–671. doi:10.1177/0013164409355699
- Løkse, M., Låg, T., Solberg, M., Andreassen, H. N., and Stenersen, M. (2017). *Teaching Information Literacy in Higher Education: Effective Teaching and Active Learning*. Cambridge, MA: Chandos Publishing.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005). Academic Self-Concept, Interest, Grades, and Standardized Test Scores: Reciprocal Effects Models of Causal Ordering. *Child. Dev.* 76 (2), 397–416. doi:10.1111/j.1467-8624.2005.00853.x
- Nierenberg, E., Låg, T., and Dahl, T. I. (2021). Knowing and Doing. *Jil.* 15 (2), 78. doi:10.11645/15.2.2795
- Oatley, K., Keltner, D., and Jenkins, J. M. (2019). *Understanding Emotions*. Hoboken, NJ: John Wiley & Sons.
- Panksepp, J. (2003). Neuroscience. Feeling the Pain of Social Loss. *Science.* 302 (5643), 237–239. doi:10.1126/science.1091062
- Pintrich, P. R., Smith, D. A. F., Garcia, T., and McKeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educ. Psychol. Meas.* 53 (3), 801–813. doi:10.1177/0013164493053003024
- Rakoczy, K., Buff, A., and Lipowsky, F. (2005). “Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie,” in *Unterrichtsqualität, Lernverhalten und mathematisches Verständnis. 1. Befragungsinstrumente* (Frankfurt, Main: GFPF ua).
- Renninger, K. A., and Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educ. Psychol.* 46 (3), 168–184. doi:10.1080/00461520.2011.587723
- Renninger, K. A., and Hidi, S. (2016). *The Power of Interest for Motivation and Engagement*. New York: Routledge.
- Rotgans, J. I. (2015). Validation Study of a General Subject-Matter Interest Measure: The Individual Interest Questionnaire (IIQ). *Health Professions Education.* 1 (1), 67–75. doi:10.1016/j.hpe.2015.11.009
- Scherer, K. R. (2005). What Are Emotions? and How Can They Be Measured?. *Soc. Sci. Inf.* 44 (4), 695–729. doi:10.1177/0539018405058216
- Schiefele, U. (1999). Interest and Learning from Text. *Scientific Stud. Reading.* 3 (3), 257–279. doi:10.1207/s1532799xssr0303\_4
- Schiefele, U., and Krapp, A. (1996). Topic Interest and Free Recall of Expository Text. *Learn. Individual Differences.* 8 (2), 141–160. doi:10.1016/S1041-6080(96)90030-8
- Schiefele, U., Krapp, A., and Winteler, A. (1992). “Interest as a Predictor of Academic Achievement: A Meta-Analysis of Research,” in *The Role of Interest in Learning and Development*. Editors K.A. Renninger, S. Hidi, and A. Krapp (Mahwah, NJ, USA: Lawrence Erlbaum Associates), 183–212.
- Schiefele, U., Krapp, A., and Winteler, A. (1988). Conceptualization and Measurement of Interest. In *Annual Meeting of the American Educational Research Association*. New Orleans, LA, USA.
- Schraw, G., Bruning, R., and Svoboda, C. (1995). Sources of Situational Interest. *J. Reading Behav.* 27 (1), 1–17. doi:10.1080/10862969509547866
- Sherry, A. (2006). Discriminant Analysis in Counseling Psychology Research. *Couns. Psychol.* 34 (5), 661–683. doi:10.1177/0011000006287103
- Silvia, P. J. (2006). *Exploring the Psychology of Interest*. Oxford: Oxford University Press.

- Silvia, P. J. (2010). Confusion and Interest: The Role of Knowledge Emotions in Aesthetic Experience. *Psychol. Aesthetics, Creativity, Arts*. 4 (2), 75–80. doi:10.1037/a0017081
- Thoman, D. B., Smith, J. L., and Silvia, P. J. (2011). The Resource Replenishment Function of Interest. *Soc. Psychol. Personal. Sci.* 2 (6), 592–599. doi:10.1177/1948550611402521
- Tracey, T. J. G. (2002). Development of Interests and Competency Beliefs: A 1-year Longitudinal Study of Fifth- to Eighth-Grade Students Using the ICA-R and Structural Equation Modeling. *J. Couns. Psychol.* 49 (2), 148–163. doi:10.1037/0022-0167.49.2.148
- Vittersø, J., Biswas-Diener, R., and Diener, E. (2005). The Divergent Meanings of Life Satisfaction: Item Response Modeling of the Satisfaction With Life Scale in Greenland and Norway. *Soc. Indic Res.* 74 (2), 327–348. doi:10.1007/s11205-004-4644-7
- Vollmeyer, R., and Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in Einem Computersimulierten System [Motivational influences on the Acquisition and Application of Knowledge in a Simulated System]. *Z. für Pädagogische Psychol.* 12 (1), 11–23.
- Vollmeyer, R., and Rheinberg, F. (2000). Does Motivation Affect Performance via Persistence? *Learn. Instruction.* 10 (4), 293–309. doi:10.1016/S0959-4752(99)00031-6
- Walton, G., and Cleland, J. (2017). Information Literacy. *Jd.* 73 (4), 582–594. doi:10.1108/JD-04-2015-0048
- Warne, R. T. (2014). A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists. *Pract. Assess. Res. Eval.* 19 (17), 1–10. doi:10.7275/sm63-7h70
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2021 Dahl and Nierenberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

## Paper 3

# Is information literacy ability, and metacognition of that ability, related to interest, gender, or education level? A cross-sectional study of higher education students

Journal of Librarianship and  
Information Science  
1–13

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09610006211058907

journals.sagepub.com/home/lis



Ellen Nierenberg  and Tove I. Dahl

UiT The Arctic University of Norway, Norway

## Abstract

How information literate are students in higher education, and how accurate is their metacognition related to that ability? Are students' perceived needs to learn more and their level of interest in becoming information literate related to their pursuit of information literacy (IL) skill development? First-year undergraduates, master's, and PhD students ( $N = 760$ ) took an objective IL test and estimated their scores both before and after the test. IL ability, as well as students' estimation of their IL ability, increased with higher education experience and IL test experience, though also varied notably within groups. Low-performers tended to overestimate their abilities, while high-performers tended to underestimate them—both evidence of the Dunning-Kruger effect. Furthermore, gender comparisons revealed that men tended to estimate higher, and more accurate, scores than women. Finally, PhD students reported greater interest in becoming information literate than undergraduates. Although undergraduates felt a greater need to learn more, PhD students were more inclined to pursue IL growth. For both groups, interest in becoming information literate correlated far more with their likelihood to invest effort into developing IL competencies than their perceived need to know more. What implications might these findings have for how we conceptualize the teaching of IL?

## Keywords

Dunning-Kruger effect, higher education, information literacy, interest, metacognition, testing

## Introduction

On January 6, 2021, the United States Capitol was stormed by individuals who believed information that results of the US presidential election were fraudulent, despite ample, and reliable media coverage of evidence to the contrary. The dissenting information sources motivating the angry skeptics were neither reliable nor credible, yet people were uncritically convinced by this misinformation, and the result of that was stupefying.

When such events occur, many questions arise. Do we know how to evaluate the quality of information sources well enough to be credibly informed citizens? Are we able to find reliable sources and use them appropriately when we produce information? Equally important, are we aware of when our abilities to evaluate, find, and use information

are insufficient? Disquieting findings from recent research such as the following imply that the answer to these questions is a resounding *no*. 53% of American adults rely on social media as their regular source of news, with Facebook topping the list of platforms (Shearer and Mitchell, 2021); 61% of Facebook users trust misinformation that they find there (Al-Zaman, 2021); and popular “fake news” stories are shared more extensively on Facebook than top stories on the same topic from more trustworthy, mainstream media (Silverman, 2016).

---

### Corresponding author:

Ellen Nierenberg, UiT The Arctic University of Norway, Postbox 6050  
Langnes, Tromsø 9037, Norway.

Email: ellen.nierenberg@uit.no

The goal of this article is to theoretically and empirically frame issues related to students' information competencies and metacognitive awareness thereof. Specifically addressed are the competencies students in higher education (HE) have for finding, evaluating, and using information sources. At the same time, we explore the relevance of student interest in possessing these skills, their felt need to develop them further, and their effort intentions related to that.

In order to examine these questions, the article begins with a presentation of the interrelated concepts of information literacy, metacognition, the Dunning-Kruger effect, and interest. Empirical evidence addressing the research questions is then analyzed and discussed. Are students metacognitively aware of their information literacy levels, and are there differences between high- and low-performers? Are there links between these initial findings and factors that could influence their motivation to learn more about information literacy?

### Information literacy

In order to best help students responsibly engage in the information-based work at higher levels of education, as well as in daily life, librarians, and teaching staff in colleges and universities have long endeavored to teach students to recognize when they have a need for information, to find and critically evaluate information sources, and to ethically and legally employ these sources in their writing. Such knowledge, skills, and dispositions are part of a diverse set of competencies known as information literacy (IL).

The definition and scope of IL has gradually evolved as technology has advanced and the creation and dissemination of information has changed character. An analysis of various definitions and frameworks (e.g. Association of College and Research Libraries [ACRL] 2000, 2015; Bruce et al., 2006; Bundy, 2004; Coonan and Secker, 2011; Secker, 2018) shows that in most, the IL-construct includes a set of competencies needed to *find*, *evaluate*, and *use* information (Nierenberg et al., 2021). These three source-related facets of IL have been extensively examined over the past 30 years, also in studies designated as for example "digital literacy" (e.g. Eshet, 2004), "media literacy" (e.g. Negi, 2018), "information problem solving" (e.g. Brand-Gruwel et al., 2009), "new literacies" (e.g. Leu et al., 2004), "multiple document comprehension" (e.g. Rouet and Britt, 2011), "civic online reasoning" (e.g. McGrew et al., 2018), and metaliteracy (e.g. Mackey and Jacobson, 2011).

With an abundance of information sources and access methods, individuals are continually faced with information choices. The abilities to find, evaluate, and use information are important competencies when navigating these choices; they are crucial both when processing and creating information and for tempering our susceptibility to misinformation. Inherent in IL is the ability to think critically, enabling us to find and make reasoned judgments

about the stream of (mis)information from its myriad of sources, whether online or offline. IL also involves the ability to cite these sources correctly when we create information in order to signal credibility and avoid plagiarism or other misuse. Effective information seeking, another core component of IL, enables us to find reliable and relevant sources. For these reasons, IL is claimed to be a vital competency—not only in education, but also in the workplace, in everyday life, in health, and for citizenship (Secker, 2018).

Because IL is multi-faceted construct with varying definitions, there is little consensus on a standard measure for assessing IL levels. The choice of assessment tools depends on not only the chosen definition, but also on the context of the research. The current study uses the definition of IL provided by Nierenberg et al. (2021: 79): "Information literacy encompasses the knowledge, skills and attitudes needed to be able to discover, evaluate and use information sources effectively and appropriately in order to answer questions, solve problems, create knowledge and learn." In terms of context, the study focuses on HE students' knowledge and self-assessments regarding finding, evaluating, and using information. Related to this, we hypothesize, is how interested a student is in being or becoming an information literate person, as well as the degree to which they feel a need to learn more IL skills and their actual intention to do so.

Finding, evaluating and using information are important competencies, and many IL researchers and practitioners recognize a need to better assess these aspects of IL (e.g. Secker and Coonan, 2011). Interestingly, most IL assessment studies rely on measures that have not been controlled for reliability and validity (Mahmood, 2017a; Walsh, 2009). Even for those that *are*, the tools' reliability is often measured with Cronbach's alpha (internal consistency; Mahmood, 2017b), which is not appropriate for multidimensional constructs such as IL (Nierenberg et al., 2021). The assessment tools employed in the current study, described in more detail in the Methods section, have evidence of both reliability and validity. With these tools, this study attempts to assess IL in HE students and to determine students' calibration accuracy.

### Metacognition

Metacognition, the overall ability to reflect upon our own thoughts, includes the ability to assess what we know (Double et al., 2018). Metacomprehension, our ability to accurately judge what we have read, is a particular aspect of metacognition (Dunlosky and Lipko, 2007; Glenberg and Epstein, 1985). In a seminal study of metacognitive awareness, Kruger and Dunning (1999) state that metacognition provides people with ". . .the ability to know how well one is performing, when one is likely to be accurate in judgment, and when one is likely to be in error" (p. 1121).

Both metacognition and metacomprehension are therefore important for effective learning, however, they can be faulty (Kruger and Dunning, 1999).

Believing that we understand something, when we, in fact, do not, can (and commonly does) lead to serious errors (Martinez, 2006; Mok et al., 2006; Norman et al., 2019). Low performers often lack these metacognitive skills (Kruger and Dunning, 1999: 1122). This miscalibration becomes problematic by limiting awareness of and attention to important learning needs.

Indeed, students' understanding of their abilities often differs from the abilities they actually exhibit in a college environment (Salisbury and Karasmanis, 2011), and new students tend to overestimate their IL competencies (Guise et al., 2008; Nierenberg and Fjeldbu, 2015; Øvern, 2018). In one study that touches on the relationship between metacognition and information seeking, students evaluated their own skills at finding information online to be adequate, even before beginning higher education, and therefore seldom sought librarian assistance (Guise et al., 2008). However, the study also found that when students experienced discrepancies between their perceived skills and their actual skills, they quickly lost confidence.

### Dunning-Kruger effect

Individuals with low abilities tend to *overestimate* their abilities and those with high abilities tend to *underestimate* their abilities (Kruger and Dunning, 1999). These biases are now referred to as the Dunning-Kruger effect. Kruger and Dunning postulated that the difference in miscalibration between *low*- and *high*-performers likely stems from erroneous judgments about *oneself* and beliefs about *others*, respectively. Researchers in multiple disciplines, including IL, have replicated Kruger and Dunning's results (Mahmood, 2016).

Händel and Fritzsche (2016) studied undergraduate students' ability to accurately estimate their performance (metacognitive monitoring ability) by asking how confident they were about their answers on a test. Low-performers were found to be "unskilled and unaware," as Kruger and Dunning argued in their 1999 article. Low-performers both overestimated their abilities, and were less accurate in their estimates than high-performers (Händel and Fritzsche, 2016). However, the low-performers were less confident in their actual test answers than high-performers. This could possibly *increase* low-performers' motivation to invest in needed learning, despite their overall overconfidence in ability, which could conceivably *decrease* their motivation. From a motivational perspective, then, the low-performing students might *not* be "doomed to remain unaware," despite their poorer performance estimation accuracy (Händel and Fritzsche, 2016).

People falling victim to the Dunning-Kruger effect regarding misinformation, and those who are unmotivated

to attempt to evaluate information otherwise, are at a disadvantage (Rosman et al., 2015). Their convictions can lead to dangerous outcomes, for example in the current pandemic. Belief in misinformation about vaccines can result in increased COVID-19 deaths, as this false information may be trusted and shared, seemingly uncritically, among misinformed individuals (Bangani, 2021).

In a larger, systematic review of empirical IL studies assessing peoples' estimated and actual IL abilities, Mahmood (2016) claims to have found convincing evidence of the Dunning-Kruger effect in 92% of relevant studies. Mahmood bases this statistic on the "overall evidence of inconsistency" between estimated and actual ability (p. 205). Mahmood suggests that these claims can be more convincingly nuanced by comparing different HE-levels, and low- and high-performing groups. Mahmood, however, does not specify whether the studies included in his review measured students' self-assessments before or after taking an objective IL test, an important distinction when interpreting results. Test-taking provides respondents with a better basis for judging their own competencies, as they then have a better idea of what the topic—in this case IL—involves (Rosman et al., 2015). However, when examining the order of testing in five randomly chosen studies from Mahmood's review, two did not specify the order of the measurements (Jackson, 2013; Oliver, 2008), two measured students' self-assessments prior to testing (Tepe and Tepe, 2015; Vickery and Cooper, 2003), and one measured their self-assessments both before and after testing (Gross and Latham, 2012).

Reliable assessment tools may be valuable for student learning of IL—including information seeking, evaluation, and use—because they can be used to increase students' metacognitive awareness of their abilities and needs. Such tools are also useful in the teaching of IL, allowing IL practitioners to more accurately assess students' competencies, as self-assessments without (or prior to) testing are particularly prone to the biases described by the Dunning-Kruger effect (Rosman et al., 2015).

### Interest

Though students may become more aware of their level of IL competence through testing, is accurately identifying a need to know more enough to motivate actual learning? Research on interest suggests that need paired with interest may be more motivating than just the identification of need alone (Rotgans and Schmidt, 2014; Tapola et al., 2013). In addition, recent research by Smarandache et al. (2021) has shown that students devote more time and effort to learning when they find the material interesting, and that this results in deeper learning. This is an important part of learning in a society with a multitude of information sources. Generating interest can therefore be fundamental for sparking a desire to take advantage of provided



**Table 1.** Research questions and hypotheses.

Research questions	Hypotheses
RQ1: How are gender and HE level (undergraduate, master's, PhD) associated with students' actual and estimated scores on an objective IL test designed to measure information seeking, evaluation, and use?	H1: Gender is not related to either actual or estimated scores. H2: The higher the education level, the higher the actual scores. H3: Scores are overestimated by undergraduates and underestimated by PhD students, before taking the test. H4: The difference between actual and estimated scores decreases as education level increases.
RQ2: Within each level of HE experience, is there a difference in how accurately low- and high-performers estimate their performance before and after taking the IL test?	H5: For students at all HE-levels, scores are overestimated by low-performers, and underestimated by high-performers, before taking the test. H6: Students at all levels estimate their scores more accurately after taking the test, especially higher-level students and high performers.
RQ3: How do students compare in their interest in being and becoming information literate and their felt need to grow in that area, and how likely do they imagine expending effort in that pursuit?	H7: Higher-level students and high performers have more interest in being or becoming information literate than others. H8: Interest in being or becoming information literate is positively related to students' felt need to learn more and their intention to put effort into IL learning. H9: The greater the felt need to learn, the greater the intention to put effort into learning.

opportunities to develop their IL skills, and the drive to continue IL learning on their own.

According to interest researchers Renninger and Hidi (2019), interest increases students' motivation to work with subject matter over time, as it has beneficial effects on both attention, memory and engagement. They have found that interest leads to meaningful learning and increased cognitive performance, as individuals with interest are more likely to expend effort in learning, developing, and employing strategies to realize their learning goals. Interest also enables the integration of information with previously acquired knowledge, allowing the formation of connections across disparate information sources (Hidi and Renninger, 2006; Kintsch, 1980)—important work for IL competence.

Knowing the difference between the nature of less and more developed interest could matter for how the teaching of IL skills is approached. Our understanding of student interest therefore has relevance for teaching and learning in general, and is likely useful for IL learning in particular. In this study, then, we explore connections between IL interest and its relationship with student awareness of their IL competence, their felt need to learn more, and their intention to actually pursue that.

**Research questions and hypotheses.** In order to determine how metacognitively aware students are of their IL levels, and to find possible links between these findings and variables that may influence their motivation to develop their IL further, we formulated three main research questions and nine corresponding hypotheses (see Table 1). The variables we chose to measure were interest, gender, and education level. As far as we know there are no empirical studies linking interest to IL, and we believe that this previously unexplored link is relevant, as interest is a factor

that increases motivation and learning. Gender is an important variable to explore in all educational contexts, and no recent studies of IL and metacognition focus on this variable. Education level can have an association with both actual and estimated scores, and is therefore also relevant to explore.

## Methods

### Participants

Data were collected from three student groups representing three levels of higher education: undergraduate, master's, and PhD (see Table 2). The undergraduate sample consisted of students in the first semester of their current study program, recruited from three universities and one university of applied sciences (UAS) across Norway. These institutions are small to medium-sized (5200–18,000 students), and offer a range of academic disciplines and professional studies. Participants came from various fields of study, including psychology, teacher education, biology, and history. Students were recruited via their learning management system (LMS) or in the classroom, and offered a small reward for their participation.

Master's and PhD students from universities in nine countries, including Norway, participated in the study. Representing a wide variety of disciplines, these students were recruited via social media (Facebook, Twitter, and Reddit) and e-mail. International graduate students were recruited when the survey was translated from Norwegian to English. Students could choose between the two language versions.

The distribution of genders for the undergraduate and master's samples is similar to that of Norway as a whole, where 66.8% and 58.4% of students are female, respectively.

**Table 2.** Descriptive statistics for three student samples.

	n	Gender			Language	
		Male	Female	Other	Norwegian	English
Undergraduate	330	123	205	2	330	0
Master's	196	76	116	4	80	116
PhD	234	78	150	6	92	142
Total	760	277	471	12	502	258

**Table 3.** ANZIL learning outcomes (Bundy, 2004) for three core facets of IL covered by the IL test.

Finding information	Evaluating information	Using information
Learning outcomes: The information literate person. . .		
Selects the most appropriate methods or tools for finding info.	Assesses the usefulness and relevance of the info. obtained	Records info. and its sources
Constructs and implements effective search strategies	Defines and applies criteria for evaluating info.	Organizes (orders/classifies/stores) info.
Obtains info. using appropriate methods	Reflects on the info. seeking process and revises search strategies as necessary	Conforms with conventions and etiquette related to access to, and use of, info.
Keeps up to date with info. sources, info. technologies, info. access tools, and investigative methods		Legally obtains, stores, and disseminates text, data, images, or sounds

In the PhD sample however, females are overrepresented by 11% compared to national levels, where 52.8% are female (Database for Statistics on Higher Education, 2019).

## Materials

Data for this study were collected with an online survey containing a self-created, psychometrically robust IL knowledge test (available at our IL assessment tools website). The test consists of 21 multiple-choice items, seven for each of the three source-focused, core components of IL described in its various definitions and frameworks: finding, evaluating, and using information. For a detailed description of test development, including evidence of reliability and validity, see Nierenberg et al. (2021). Items were designed to test learning outcomes from the *Australian and New Zealand Information Literacy Framework* (ANZIL; Bundy, 2004) for the three aforementioned facets (see Table 3). The ANZIL framework is based on the previous gold standard for assessing IL, the *Information Literacy Competency Standards for Higher Education* (ACRL, 2000). The ACRL Standards, which guided instruction and facilitated assessment of IL by providing explicit learning outcomes, were replaced in 2015 by the more conceptual *Framework for Information Literacy for Higher Education* (ACRL, 2015). The ANZIL and ACRL standards do not conflict with more recent frameworks, but detail concrete learning outcomes that aid in teaching and measuring students' developmental journeys toward greater IL, with its other facets, as well.

Each item in the IL knowledge test has four randomly ordered alternative answers, one of which is *most correct*. There is no option for "I do not know," and each item must be answered in order to proceed to the next. This example item, with the correct answer in italics, relates to the use of sources: "What is the most important reason to use sources when writing a paper? (a) *To support arguments*; (b) To avoid plagiarism; (c) To show that you've read the sources; (d) To satisfy the requirements of the assignment."

Several methods were used to evaluate the IL knowledge test for validity and reliability, as described in Nierenberg et al. (2021). Expert evaluations and think-aloud protocols ensured the validity of test items. After four IL experts evaluated the items for their clarity, content accuracy, and objectivity, think-aloud protocols with students in the target group were held to check items for readability and comprehension. The IL test demonstrated high test-retest reliability, a better measure of reliability than internal consistency for tests/indexes of multidimensional constructs such as IL, as argued in the aforementioned article. For a more detailed description of the development of the IL knowledge test, including item selection criteria and evidence for its reliability and validity, see Nierenberg et al. (2021).

## Procedure

In order to measure participants' metacognition of their IL knowledge, all respondents estimated how many of the

questions they believed they had answered correctly, *after* having taken the test. After data collection had begun, we realized that it would be useful to also have students estimate their scores *prior* to taking the IL test, based on the working definition of IL provided at the beginning of the survey. A total of 308 respondents (70 undergraduates, 122 master's students, 116 PhD students) were therefore given the opportunity in a revised version of the questionnaire to also estimate, before the test, how many of the 21 questions they predicted they would answer correctly. Additionally, 70 undergraduate and 64 PhD students were also asked to rate on a scale from 1 (not at all true for me) to 6 (very true for me) two additional statements before the test: "I am interested in being or becoming an information literate person" (IL interest) and "I need more skills in information literacy" (IL need). After taking the test and receiving their final score, they rated the same two statements again, as well as a third question, "Knowing myself, I will make the effort to develop stronger information literacy skills" (IL effort; 1 = not likely at all, 6 = highly likely).

Data were collected from August 2019 to December 2020. Test results were analyzed using IBM SPSS Statistics 26 (2019).

To determine whether there are differences within each educational level for students scoring low and high on the IL test, results for each level were split into two subgroups—those scoring below the median and those scoring above the median (the median being different for each HE level).

## Results

Results are arranged in the order of the three research questions.

### *RQ1. IL levels and score estimation accuracies by gender and HE level*

Table 4 provides descriptive statistics for IL test scores for the three samples, by gender. The maximum possible score on the IL test is 21 points.

**Gender and HE level comparisons.** Table 5 shows actual IL test scores and the difference between actual and estimated scores (estimation accuracy scores) before and after the test. Positive estimation accuracy mean values indicate overestimated scores, while negative mean values indicate underestimated scores. The closer the value to zero, the better the student's metacognition. Students scoring below or above the median score were divided into low- and high-performing groups for each HE level (For additional descriptive statistics and analyses of students' estimated scores, see Supplemental Material 1, available at our IL assessment tools website, <https://site.uit.no/troils/>).

**Table 4.** Descriptive statistics for IL test scores, by gender and HE level.

Level	n	IL test score			SD	95% CI	
		M	Min.	Max.		LL	UL
<b>Undergraduates</b>							
All	330	12.71	3	20	3.12	12.37	13.04
Male	123	13.00	3	20	3.10	12.45	13.55
Female	205	12.51	3	20	3.10	12.09	12.94
Other	2	14.50	10	19	6.36	—	—
<b>Master's</b>							
All	196	16.06	8	21	2.49	15.71	16.41
Male	76	15.67	8	21	2.57	15.08	16.26
Female	116	16.20	10	20	2.38	15.76	16.64
Other	4	19.25	19	20	0.50	18.45	20.05
<b>PhD</b>							
All	234	16.89	10	21	2.13	16.62	17.17
Male	78	16.90	10	21	2.39	16.36	17.44
Female	150	16.84	10	21	2.00	16.52	17.16
Other	6	18.17	16	20	1.47	16.62	19.71

CI: confidence interval; LL: lower limit; UL: upper limit.

Data from the English and Norwegian versions of the IL test are combined in this table and in subsequent analyses since no statistically significant differences were found between participant performances on the two language versions (Nierenberg et al., 2021).

**IL test score comparison.** When excluding the small gender group "other," a two-way ANOVA of IL test scores by HE level (undergraduate, master's, and PhD) and gender (male, female), showed a main effect for HE level,  $F(2, 745) = 187.47$ ,  $p < 0.001$ ,  $\eta^2 = 0.34$ . Mean IL test scores increased with HE level, as could be expected. Bonferroni post hoc tests showed that each successive HE level scored significantly higher than the level below it ( $p < 0.01$  for all comparisons). There was no significant difference between men and women on overall test scores. No main effect was found for gender, and the interaction between gender and HE level was not statistically significant.

**Estimation accuracy comparison.** Student estimation accuracy was calculated by subtracting each individual's actual IL test score from their estimated IL test score, both before and after taking the test. Figure 1 illustrates overall estimation accuracy scores for those students who made test score estimates. Overestimated scores are on the right side of zero and the underestimated scores are on the left. The closer the value to zero, the more accurate the estimate.

Accuracy differences by HE level and gender (male and female) were tested with two-way ANOVAs for before- and after-test accuracy scores. Effect size was measured with eta squared ( $\eta^2$ ). A two-way ANOVA of *before-test* accuracy scores by HE level and gender showed a main effect for HE level,  $F(2, 297) = 3.82$ ,  $p < 0.05$ ,  $\eta^2 = 0.03$ , but not for gender. Bonferroni post hoc tests showed that undergraduate and master's student estimation accuracies were not significantly

**Table 5.** Descriptive statistics for IL test scores and estimation accuracy scores, before and after IL test, by HE level, performance group, and gender.

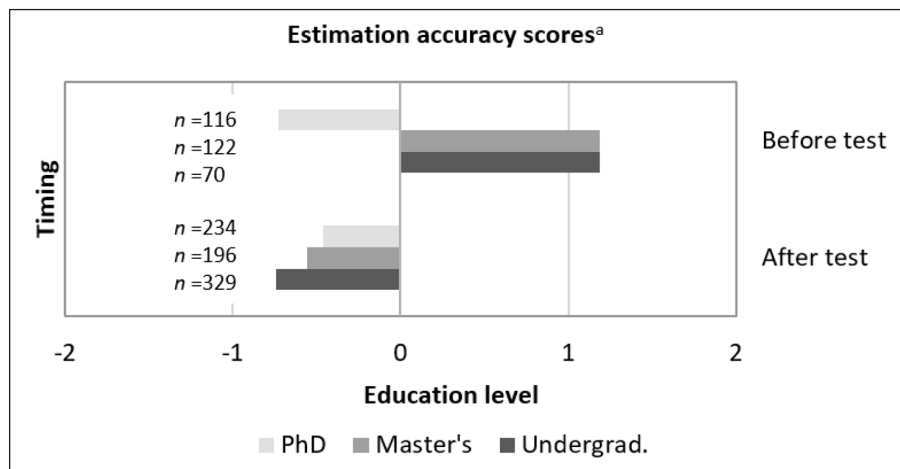
Group (median)	Test score <sup>a</sup>			Estimation accuracy score <sup>b</sup>											
	M	SD	n	95% CI <sup>c</sup>		Before <sup>d</sup>			95% CI		After			95% CI	
				LL <sup>e</sup>	UL <sup>f</sup>	M	SD	n	LL	UL	M	SD	n	LL	UL
Undergrad. (13)	<b>12.71</b>	<b>3.12</b>	<b>330</b>	<b>12.37</b>	<b>13.04</b>	<b>1.19</b>	<b>5.27</b>	<b>70</b>	<b>-0.07</b>	<b>2.44</b>	<b>-0.74</b>	<b>4.06</b>	<b>329</b>	<b>-1.24</b>	<b>-0.24</b>
Low <sup>g</sup>	9.72	2.11	137	9.37	10.08	4.57	5.28	21	2.17	6.97	0.48	4.83	137	-0.34	1.30
High <sup>h</sup>	15.43	1.44	145	15.19	15.66	-0.76	4.27	41	-0.34	1.30	-1.94	4.09	144	-2.62	-1.27
Master's (17)	<b>16.06</b>	<b>2.49</b>	<b>196</b>	<b>15.71</b>	<b>16.41</b>	<b>1.22</b>	<b>4.15</b>	<b>122</b>	<b>0.45</b>	<b>1.93</b>	<b>-0.56</b>	<b>3.59</b>	<b>196</b>	<b>-1.06</b>	<b>-0.05</b>
Low	13.99	1.79	96	13.63	14.35	2.75	4.60	64	1.59	3.88	0.90	3.70	96	0.15	1.65
High	18.65	0.74	63	18.46	18.84	-0.62	2.72	40	-1.49	0.24	-2.16	2.95	63	-2.90	-1.42
PhD (17)	<b>16.89</b>	<b>2.13</b>	<b>234</b>	<b>16.62</b>	<b>17.17</b>	<b>-0.73</b>	<b>4.43</b>	<b>116</b>	<b>-1.55</b>	<b>0.08</b>	<b>-0.46</b>	<b>3.46</b>	<b>234</b>	<b>-0.91</b>	<b>-0.02</b>
Low	14.54	1.52	80	14.20	14.88	1.86	4.40	37	0.40	3.33	1.01	4.18	80	0.08	1.94
High	18.85	0.9	93	18.66	19.03	-2.38	3.50	52	-3.36	-1.41	-1.62	2.52	93	-2.14	-1.10
Gender															
Male	<b>14.83</b>	<b>3.24</b>	<b>277</b>	<b>14.45</b>	<b>15.21</b>	<b>0.90</b>	<b>4.80</b>	<b>117</b>	<b>0.02</b>	<b>1.78</b>	<b>0.13</b>	<b>3.93</b>	<b>276</b>	<b>0.60</b>	<b>-0.16</b>
Female	<b>14.80</b>	<b>3.31</b>	<b>471</b>	<b>14.50</b>	<b>15.10</b>	<b>0.22</b>	<b>4.51</b>	<b>186</b>	<b>-0.44</b>	<b>0.87</b>	<b>-1.01</b>	<b>4.01</b>	<b>471</b>	<b>-1.37</b>	<b>-0.65</b>

Data from the English and Norwegian versions of the IL test are combined also in this table. Scores for the different levels of HE overall and gender are highlighted in bold lettering.

<sup>a</sup>For all respondents; <sup>b</sup>Estimation accuracy score is the estimated score minus the actual score of those who estimated; <sup>c</sup>CI: confidence interval;

<sup>d</sup>Although all respondents estimated their scores *after* the IL test, a smaller subset of students estimated their scores *before* the test; <sup>e</sup>lower limit;

<sup>f</sup>upper limit; <sup>g</sup>Below median group; <sup>h</sup>Above median group.

**Figure 1.** Estimation accuracy scores, before and after test.

Positive values indicate overestimation of actual scores; negative values indicate underestimation of actual scores.

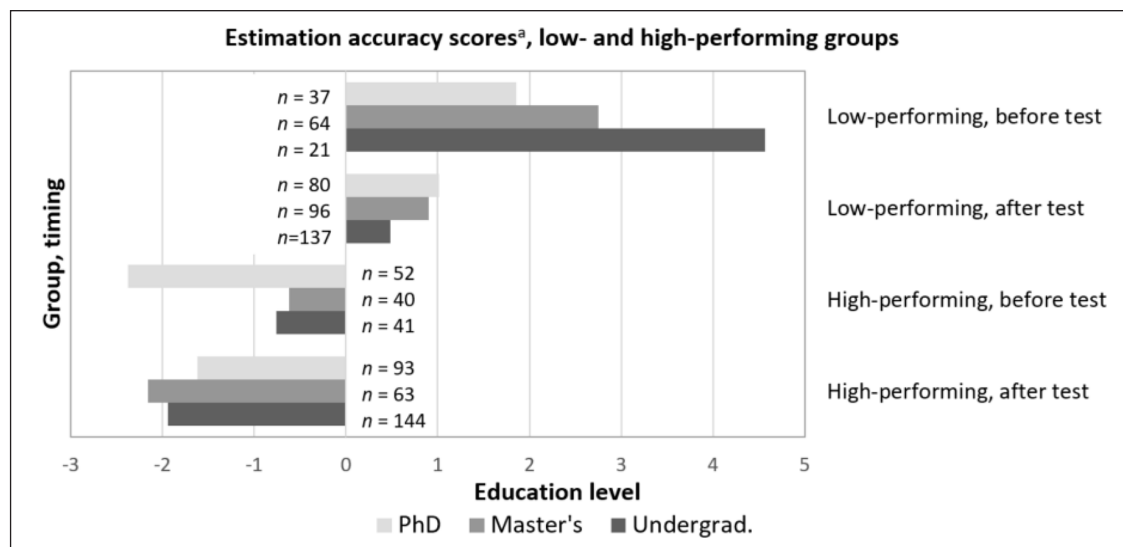
<sup>a</sup>Estimation accuracy score = estimated score minus actual score.

different, while PhD student estimation accuracy was significantly higher than undergraduate ( $p < 0.01$ ) and master's student estimation accuracy ( $p < 0.05$ ).

A two-way ANOVA of *after-test* accuracy scores by HE level and gender showed no main effect for HE level, but a main effect for gender,  $F(1, 741) = 10.44, p = 0.001, \eta^2 = 0.01$ . The HE level by gender interaction was statistically significant,  $F(2, 741) = 4.38, p < 0.05, \eta^2 = 0.01$ . Three independent sample *t*-tests comparing men's and women's scores—with effect size measured with Cohen's *d*—showed that undergraduate men made significantly more accurate estimations than undergraduate women,  $t(325) = 4.04, p < 0.001, d = 0.15$ .

There were no significant gender differences found for the master's and PhD students.

Together, these findings show that students at higher HE levels have higher estimates of their IL competency, and that they are better at accurately estimating their IL test scores compared to students at lower HE levels. However, once they have had a chance to calibrate their understanding with the IL test itself, that difference more or less disappears among all HE levels. Furthermore, though men make more accurate after-test estimates than women at the undergraduate level, women's estimates become equally accurate with more HE experience.



**Figure 2.** Estimation accuracy scores, low- and high-performing groups, before and after IL test.

Positive values indicate overestimation of actual scores; negative values indicate underestimation of actual scores.

<sup>a</sup>Estimation accuracy score = estimated score minus actual score.

### RQ2. Score estimation accuracies: Comparison of low- and high-performing groups by HE level

There may be more difference within HE levels between low- and high-performers on the IL test than these overall analyses capture, particularly in terms of how much students under- or overestimate their IL test score—both before and after taking the IL test—and in terms of their desire and commitment to developing those competencies. To test this, students were divided into a low- or high-performing group by median split for each of the three HE levels (the median varied by HE level as noted in Table 5). Students who scored below the median in each HE level were put in a *low-performing* group, and those who scored above the median in each HE level were put in a *high-performing* group. Estimation accuracy scores for the low- and high-performing groups at different HE levels are illustrated in Figure 2.

As for estimation accuracy scores, two-way ANOVAs by HE level and performance level group indicated a main effect for HE level in *before-test* accuracy,  $F(1, 254) = 5.05$ ,  $p < 0.01$ ,  $\eta^2 = 0.04$ , but not *after-test* accuracy. For *before-test* accuracy, a Bonferroni post hoc test shows that there were significant differences between master's and PhD students,  $p < 0.005$ , and between undergraduate and PhD students,  $p < 0.05$ . The difference between undergraduate and master's students however was not significant. A main effect was also found for performance level group. Differences between low- and high-performing groups are significant both before,  $F(1, 254) = 62.04$ ,  $p < 0.001$ ,  $\eta^2 = 0.20$ , and after,  $F(1, 612) = 66.87$ ,  $p < 0.001$ ,  $\eta^2 = 0.10$ , taking the test. This indicates that there are significant differences in the low and high groups' abilities to accurately estimate their scores.

Taken together, results indicate that students have a diffuse sense of their IL skill before taking the IL test, while after taking it, most become better at calibrating their score estimates. Interestingly, although lower-performing students tend to overestimate their skill rather than underestimate it like their higher-performing peers, low-performers are more accurate in making after-test estimations than high-performers. These metacognitive differences may have differential relevance for the students' felt need or intent to develop their IL competencies.

### RQ3. Interest, need, and effort: Comparisons by HE level and performance level

In order to determine whether taking the IL test influenced students' interest in becoming information literate, a selection of undergraduate ( $n = 70$ ) and PhD ( $n = 64$ ) students were asked two questions both before and after the test, and a third question after the test.

The first question ("interest") asked about the student's interest in being or becoming information literate, and the second ("need") asked whether they believe that they need more skills in order to do this. The third question ("effort") asked whether, knowing themselves, they would put the effort into obtaining additional IL knowledge and skills. Before- and after-test answers to the interest and need questions correlated highly with each other ( $p < 0.001$ ), indicating that the act of taking the test did not significantly change answers regarding interest or the perceived need to know more (see Table 6). Their likelihood to do anything about that (effort), however, was significantly correlated only with their interest in being/becoming an information literate person (both before and after the test), but not with their

**Table 6.** Descriptive statistics and correlations (Pearson's  $r$ ) for pre- and post-test questions about *interest* in being/becoming information literate, *need* for more skills, and *likely effort* students would put into learning more about IL.

Question	<i>n</i>	<i>M</i>	<i>SD</i>	1.	2.	3.	4.
1. Pre-test: IL interest	134	5.27	0.99	—			
2. Post-test: IL interest	124	5.26	0.85	0.517**	—		
3. Pre-test: IL need	134	4.46	1.17	0.256**	0.183*	—	
4. Post-test: IL need	124	4.85	1.00	0.141	0.325**	0.547**	—
5. Post-test: IL effort	124	4.39	1.08	0.440**	0.478**	0.119	0.153

\*Correlation is significant at the 0.05 level (two-tailed).

\*\*Correlation is significant at the 0.01 level (two-tailed).

felt need for more skill. Furthermore, their perceived need to grow, and what they believed they would actually do about that, were hardly correlated at all (For additional descriptive statistics and analyses related to interest, need and effort, see Supplemental Material 2, available at our IL assessment tools website, <https://site.uit.no/троils/>).

A two-way repeated measure MANOVA was performed with HE group (undergraduate, PhD) by performance level group (low, high), comparing the two dependent variables—student IL interest and perceived skill need—before and after the test. Between groups, a main effect was found for HE level,  $F(2, 101)=9.41$ ,  $p<0.001$ ,  $\eta^2=0.16$ , and performance level,  $F(2, 101)=4.05$ ,  $p<0.05$ ,  $\eta^2=0.07$ . PhD students reported significantly greater interest both before and after the test, than undergraduates. Nevertheless, undergraduates reported significantly greater need for skills both before and after the test than PhD students.

Within groups, an interaction was found between time (before and after the test) and performance level on interest in being/becoming an information literate person,  $F(2,101)=3.21$ ,  $p<0.05$ ,  $\eta^2=0.06$ . While the high-performing undergraduates were more interested both before and after the test than their lower-performing peers, there was a tendency for higher-performing PhD students to be less interested before the test than their lower-performing peers, though more interested than their lower-performing peers after the test ( $p=0.057$ ).

In sum, between groups, PhD students reported a greater interest in being or becoming information literate than the undergraduates, while undergraduates reported recognizing a greater need for more IL skill than the PhD students. Within groups, the high-performing undergraduate student interest was greater than that of their lower-performing peers and that relationship was relatively stable across time. Meanwhile, though the lower-performing PhD students were more interested than their high-performing peers before the test, that relationship switched after the test.

*Likelihood to make the effort to develop stronger information literacy skills.* After taking the test, the same selection of undergraduates and PhD students rated the likelihood that they would make an effort to develop stronger IL skills. Although a two-way ANOVA by HE level and

performance level indicated no significant difference between low- and high-performers, it did show a main effect for HE level,  $F(1, 105)=12.21$ ,  $p=0.001$ ,  $\eta^2=0.11$ . PhD students reported being more likely to make the effort to develop their IL skills than undergraduates were.

The correlation between interest in being or becoming information literate and the likelihood of making an effort to develop IL skills is significant for both HE levels tested, and is higher for PhD students ( $r=0.51$ ) than for undergraduates ( $r=0.42$ ).

## Discussion

A main goal of this study was to determine whether levels of IL, based on results of an objective 21-item IL test, correspond to student metacognitive awareness of that ability. Unlike similar research, though, we specifically analyzed associations to students' gender, HE level, and performance level. Furthermore, we investigated whether students' IL performance is related to (1) their interest in being or becoming an information literate person, (2) their perceived need for increasing their proficiency in IL, and (3) how likely they imagine putting effort into pursuing that need.

Our study, as far as we know, is the first to address gender differences in self-assessments of IL ability and their accuracy. We found that although men and women had nearly identical mean test scores, as predicted in H1, men's estimated scores were both higher and more accurate than women's after the test—an unexpected result. These differences were statistically significant for undergraduates' and master's students' after-test estimates and accuracy. Though several other IL studies have shown that men tend to have higher self-efficacy than women (e.g. Punter et al., 2017; Taylor and Dalal, 2017), our results suggest that men additionally tend to have better IL metacognition.

Overall, we found that student IL knowledge levels generally improve with HE experience, as hypothesized in H2. When estimating their test scores, students at all HE levels vary both in the estimates and in their accuracy, which gives an indication of their metacognitive awareness. Before the test, undergraduate and master's students tended to overestimate their skills, while PhD students

tended to underestimate their skills. These findings are in accordance with H3. In addition, PhD students had higher estimation accuracies than lower level students, as predicted in H4. After taking the test, all HE levels underestimated their skills, though their estimation accuracy improved (see Figure 1). One could even say that their estimation accuracy, and thereby their metacognitive monitoring ability, was remarkably good. H6 predicted this higher accuracy in scores estimated after the test, but we had not foreseen that all HE levels would underestimate their actual scores after the test. Nevertheless, the higher the HE level, the more accurate the after-test estimate, as other IL researchers also have found (Mahmood, 2016). The overall improvement in calibration is most likely a result of students anchoring their estimations based on the test content; they were primed to become metacognitively aware of their ability by seeing the questions. This was an important effect of the measure itself.

Performance-level comparisons revealed that low-performers on the IL test tended to overestimate their skills, while high-performers tended to underestimate their skills, both before and after the test, as predicted in H5 and H6. In addition, the low-performers estimated their scores more accurately after taking the test than before taking the test, a healthy sign of knowledge calibration.

Our findings thereby illustrate a typical example of the Dunning-Kruger effect. The lower the student's HE level and/or performance level, the more likely they were to overestimate their abilities (Kruger and Dunning, 1999; Mahmood, 2016). Explained simply, they cannot know what they do not know. Conversely, those with the most HE experience and/or a high performance level, tended to underestimate their scores. Kruger and Dunning (1999) might attribute this miscalibration to a comparative error associated with beliefs about others' performance. In other words, they are aware of what they do not know, and believe that others may have this knowledge.

The inability to recognize incompetence has many implications. One consequence of overestimating performance, as observed by Guise et al. (2008), is that students may avoid seeking assistance from librarians or teachers, believing instead—perhaps mistakenly—that they can find reliable information sources on their own. However, without the competencies necessary to find and evaluate sources, a student may end up trusting conspiracy theories or other misinformation found with a simple Google search. Recognizing the need for improvement is essential for voluntary self-improvement (Pennycook et al., 2017), and an important motivation for learning.

Contrary to H6 and to prior research on the Dunning-Kruger effect, we did find an unexpected result. While high-performing students at all HE levels tended to underestimate their scores after the test, undergraduate and master's students' after-test estimates deviated surprisingly more from their actual scores than their before-test estimates. This may indicate that the high-performing

undergraduate and master's students, more than the PhD students, are still calibrating their understanding of their IL proficiency even after the test, though it is intriguing that they are poorer at this than their low-performing peers. Other studies show more *accurate* calibrations in after-test estimates (cf. Händel and Fritzsche, 2016).

This fits with the Gross and Latham (2007) argument, however, that more skilled individuals may adjust their estimates by considering how their performance might compare with that of other experts. Likewise, "imposter syndrome," where people fear they lack the knowledge or skills they have been ascribed, has been noted among high achieving students in other research as well (see, e.g. Craddock et al., 2011; Gibson-Beverly and Schwartz, 2008).

Meanwhile, some researchers, for example Krueger and Mueller (2002), argue that what others report as evidence of the Dunning-Kruger effect may in some cases actually be a statistical artifact, caused by regression to the mean (RTM) and/or the better-than-average effect (BTA). RTM can occur when unusually small or large measurements are followed by measurements closer to the mean. If RTM was a factor in our study, estimation accuracy scores would have been more evenly distributed around zero in Figures 1 and 2. Our distributions therefore suggest real variation as opposed to statistical artifacts. Meanwhile, BTA refers to the observation that the majority of people predict higher-than-average scores on cognitive tests, such as the IQ test, when the mean score is known. However, students in our study were unaware of the IL test group mean scores, rendering the BTA irrelevant. Finally, in research on the Dunning-Kruger effect, it is common to compare participants in the lowest and highest quartiles. By using a median split to distinguish low- and high-performers in each HE group, the statistical differences we found were therefore even more pronounced.

Finally, a unique contribution to the field is our investigation of additional factors related to student and metacognitive competence that may affect how students go about addressing gaps in their IL competencies. Indeed, as hypothesized in H7, interest in being or becoming information literate was higher among the high performers in general and among PhD students more than undergraduates, though the high-performing PhD students reported being more interested after the test than before. Interest was positively related to students' felt need to learn more and their intention to put effort into IL learning, as predicted in H8. Importantly, student interest is more correlated with their felt need to learn after they have taken the test, than it is before—though PhD students reported being significantly more likely to act on that need than the undergraduates. However, the greatest correlation was between student interest in being or becoming information literate and their likelihood to invest effort in developing IL skills—both before and after taking the test. In that sense, arousing IL interest may be more effective for inspiring action than identifying a learning need, particularly among

lower-performing students. This result deviated from H9, which predicted that the felt need to learn more about IL would have a higher correlation to students' intended effort to learn. These findings are correlational, though, and should therefore be interpreted with that in mind, providing grounds for further study of these relations.

### Limitations

In order to make bolder claims about gender comparisons, sampling is important. We strove to complement existing research with a sample that had gender distributions representative for Norwegian HE. We were successful in this for the undergraduate and master's levels, but not for the PhD level, where we had an overrepresentation of women. This limitation of the study may explain why men in the PhD group did not show metacognitive and test-estimate accuracy advantages, though more research is required to test this.

Another limitation of this study is a cultural difference between samples that may have affected results; while all undergraduate participants took the Norwegian version of IL test, one-third of graduate students took the English version. In addition, four questions were added to the survey after data collection had begun, and were therefore not answered by all respondents.

### Implications and future research

Of methodological interest, student estimates of their IL knowledge were often inaccurate, especially when predicting scores *before* taking the IL test. This implies that "cold," or before-test, predictions are not as reliable a measure of student IL competence and should be interpreted with caution also in other research. This study found that estimation accuracy improves however, when students estimate their scores *after* the test. This implies that taking a reliable IL test has value in helping students develop a metacognitive awareness of their IL ability. Taking the IL test in this study also led to an increased correlation between interest and the felt need to learn more.

Accurate score estimates are also helpful for those who teach IL, regarding *who* to help (prioritize low-performers and low HE levels), and *how* to help (testing contributes to skill awareness). Our study also shows that taking an IL test helps students calibrate their understanding of what they know. With increased metacognition, undergraduates grow most in understanding their need to learn more. In this way, in addition to affirming the value of the IL test per se, the test scores also clarify for students the relevance of IL instruction for them, personally.

This discrepancy—between what students *initially* believe about their competence and how they perceive it *after* testing—is important for IL research that bases its findings on self-reports (Gross and Latham, 2012). Conclusions from studies based on pre-test self-assessments may be less accurate than those based on post-test

self-assessments (Rosman et al., 2015). For example, despite being provided with a definition of IL upon which to base their pre-test competence judgment in this study, students' post-test judgments aligned significantly better with their actual demonstrated competence.

Low-performers may discover or reinforce a need to improve. Knowing themselves, however, acting on that is more likely associated with their interest in being or becoming an IL person than their identified need. In addition, in terms of HE level, results indicate that interest in being or becoming information literate, and the likelihood of making an effort to develop IL skills, are correlated and increase with each level. So, worthy of further investigation in future research is the role interest plays both in the HE trajectory of IL learning, and in helping students identify their skill levels.

Though helping especially lower-performing students develop interest in being or becoming IL may be valuable for improving student IL learning, it may be equally valuable to counteract student loss of interest when they do not perform well, by simultaneously developing feedback opportunities to support positive growth in IL self-efficacy (Guise et al., 2008). Interventions designed for students with low IL performance and/or interest levels with this in mind may be especially worthwhile, and a focus of future research.

### Conclusion

Differences in IL competencies and metacognition related to those competencies in students of different genders and HE levels—both understudied to date—have been explored in this article. In addition to confirming results from previous studies on the existence of the Dunning-Kruger effect in the IL domain, this study makes several new contributions to the field. Findings indicate that (1) in general, IL metacognition increases with HE level and performance level, (2) in some groups, men tend to have better metacognitive awareness of their IL competency than women, and (3) interest in being or becoming information literate is highly associated with the likelihood of putting effort into learning more, more so than need, implying that getting students interested in IL may be as relevant for IL instruction as the teaching of the competencies themselves.

It is in all our best interest to use the platform of academia to help students develop the knowledge and skills they need to become intelligent information consumers and contributors. It is relevant not only for academic performance, but also for informed social engagement, and it is especially important in today's "post-truth era," characterized by a growing disregard for facts and distrust in mainstream media.

Because of the larger, more diverse sample than that used in previous research on the Dunning-Kruger effect, this study identifies nuances in that effect related to HE experience, gender and performance level. Intriguing



connections between student interest, need, and likely effort in attending to their IL growth have also been highlighted. These findings can be used in future research and practice in order to design even better approaches to student IL development in all levels of higher education, and, just as importantly, to help counteract people's vulnerability to false or misleading information that can have truly dangerous consequences.

### Acknowledgements

The authors would like to thank bachelor student Mina Taaje Sønsterud (UiT The Arctic University of Norway) for her valuable contributions to this article.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Ellen Nierenberg  <https://orcid.org/0000-0001-8666-8092>

### Supplemental material

Supplemental material for this article is available online at <https://site.uit.no/troils/>.

### References

- Al-Zaman MS (2021) Social media and COVID-19 misinformation: How ignorant Facebook users are? *Heliyon* 7(5): e07144.
- Association of College and Research Libraries (2000) Information literacy competency standards for higher education. Available at: <http://hdl.handle.net/11213/7668> (accessed 25 February 2021).
- Association of College and Research Libraries (2015) Framework for information literacy for higher education. Available at: <https://www.ala.org/acrl/standards/ilframework> (accessed 25 February 2021).
- Bangani S (2021) The fake news wave: Academic libraries' battle against misinformation during COVID-19. *The Journal of Academic Librarianship* 47(5): 102390.
- Brand-Gruwel S, Wopereis I and Walraven A (2009) A descriptive model of information problem solving while using internet. *Computers & Education* 53(4): 1207–1217.
- Bruce C, Edwards S and Lupton M (2006) Six frames for information literacy education: A conceptual framework for interpreting the relationships between theory and practice. *Innovation in Teaching and Learning in Information and Computer Sciences* 5(1): 1–18.
- Bundy A (2004) *Australian and New Zealand Information Literacy Framework: Principles, Standards and Practice*, 2nd edn. Adelaide: Australian and New Zealand Institute for Information Literacy.
- Coonan E and Secker J (2011) A new curriculum for information literacy (ANCIL): Curriculum and supporting documents. Available at: <https://www.repository.cam.ac.uk/handle/1810/244638> (accessed 24 April 2021).
- Craddock S, Birnbaum M, Rodriguez K, et al. (2011) Doctoral students and the impostor phenomenon: Am I smart enough to be here? *Journal of Student Affairs Research and Practice* 48(4): 429–442.
- Database for Statistics on Higher Education (2019) Statistikk fordelt etter område. Available at: <https://dbh.hkdir.no/tall-og-statistikk/statistikk-meny/studenter> (accessed 29 June 2020).
- Double KS, Birney DP and Walker SA (2018) A meta-analysis and systematic review of reactivity to judgements of learning. *Memory* 26(6): 741–750.
- Dunlosky J and Lipko AR (2007) Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science* 16(4): 228–232.
- Eshet Y (2004) Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia* 13(1): 93–106.
- Gibson-Beverly G and Schwartz JP (2008) Attachment, entitlement, and the impostor phenomenon in female graduate students. *Journal of College Counseling* 11(2): 119–132.
- Glenberg AM and Epstein W (1985) Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11(4): 702–718.
- Gross M and Latham D (2007) Attaining information literacy: An investigation of the relationship between skill level, self-estimates of skill, and library anxiety. *Library & Information Science Research* 29(3): 332–353.
- Gross M and Latham D (2012) What's skill got to do with it? Information literacy skills and self-views of ability among first-year college students. *Journal of the American Society for Information Science and Technology* 63(3): 574–583.
- Guisse JL, Goosney J, Gordon S, et al. (2008) Evolution of a summer research/writing workshop for first-year university students. *New Library World* 109(5/6): 235–250.
- Händel M and Fritzsche ES (2016) Unskilled but subjectively aware: Metacognitive monitoring ability and respective awareness in low-performing students. *Memory & Cognition* 44(2): 229–241.
- Hidi S and Renninger KA (2006) The four-phase model of interest development. *Educational Psychologist* 41(2): 111–127.
- Jackson C (2013) Confidence as an indicator of research students' abilities in information literacy: A mismatch. *Journal of Information Literacy* 7(2): 149–152.
- Kintsch W (1980) Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics* 9(1–3): 87–98.
- Krueger J and Mueller RA (2002) Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology* 82(2): 180–188.
- Kruger J and Dunning D (1999) Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121–1134.
- Leu DJ, Kinzer CK, Coiro JL, et al. (2004) Toward a theory of new literacies emerging from the Internet and other information and communication technologies. In: Rudell RB and Unrau N (eds) *Theoretical Models and Processes of*

- Reading*, 5th edn. Newark, DE: International Reading Association, pp.1570–1613.
- McGrew S, Breakstone J, Ortega T, et al. (2018) Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education* 46(2): 165–193.
- Mackey TP and Jacobson TE (2011) Reframing information literacy as a metaliteracy. *College & Research Libraries* 72(1): 62–78.
- Mahmood K (2016) Do people overestimate their information literacy skills? A systematic review of empirical evidence on the Dunning-Kruger effect. *Communications in Information Literacy* 10(2): 199–213.
- Mahmood K (2017a) Reliability and validity of self-efficacy scales assessing students' information literacy skills: A systematic review. *The Electronic Library* 35(5): 1035–1051.
- Mahmood K (2017b) A systematic review of evidence on psychometric properties of information literacy tests. *Library Review* 66(6/7): 442–455.
- Martinez ME (2006) What is metacognition? *Phi Delta Kappan* 87(9): 696–699.
- Mok MMC, Lung CL, Cheng DPW, et al. (2006) Self-assessment in higher education: Experience in using a metacognitive approach in five case studies. *Assessment & Evaluation in Higher Education* 31(4): 415–433.
- Negi MU (2018) Fake news and information literacy: A case study of Doon University, Dehradun. *International Research: Journal of Library and Information Science* 8(2): 202–214.
- Nierenberg E and Fjeldbu ØG (2015) How much do first-year undergraduate students in Norway know about information literacy? *Journal of Information Literacy* 9(1): 15–33.
- Nierenberg E, Låg T and Dahl TI (2021) Knowing and doing: The development of information literacy measures to assess knowledge and practice. *Journal of Information Literacy* 15(2): 78–123.
- Norman E, Pfuhl G, Sæle RG, et al. (2019) Metacognition in psychology. *Review of General Psychology* 23(4): 403–424.
- Oliver B (2008) Commencing undergraduates' self-efficacy and ability in finding academic resources: Are they improving? *Studies in Learning Evaluation Innovation and Development* 5(3): 1–8.
- Øvern KM (2018) Information use among first-year students in health sciences: Is an intervention needed? *LIBER Quarterly* 28(1): 1–28.
- Pennycook G, Ross RM, Koehler DJ, et al. (2017) Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin & Review* 24(6): 1774–1784.
- Punter RA, Meelissen MR and Glas CA (2017) Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICILS 2013. *European Educational Research Journal* 16(6): 762–780.
- Renninger KA and Hidi SE (2019) Interest development and learning. In: Renninger KA and Hidi SE (eds) *The Cambridge Handbook of Motivation and Learning*. Cambridge: Cambridge University Press, pp.265–290.
- Rosman T, Mayer A-K and Krampen G (2015) Combining self-assessments and achievement tests in information literacy assessment: Empirical results and recommendations for practice. *Assessment & Evaluation in Higher Education* 40(5): 740–754.
- Rotgans JI and Schmidt HG (2014) Situational interest and learning: Thirst for knowledge. *Learning and Instruction* 32: 37–50.
- Rouet J-F and Britt MA (2011) Relevance processes in multiple document comprehension. In: McCrudden MT, Magliano JP and Schraw G (eds) *Text Relevance and Learning From Text*. Charlotte, NC: Information Age Publishing, pp.19–52.
- Salisbury F and Karasmanis S (2011) Are they ready? Exploring student information literacy skills in the transition from secondary to tertiary education. *Australian Academic & Research Libraries* 42(1): 43–58.
- Secker J (2018) The revised CILIP definition of information literacy. *Journal of Information Literacy* 12(1): 156–158.
- Secker J and Coonan E (2011) A new curriculum for information literacy (ANCIL) – curriculum and supporting documents. Available at: <http://www.dspace.cam.ac.uk/handle/1810/244638> (accessed 2 May 2020).
- Shearer E and Mitchell A (2021) News use across social media platforms in 2020. Available at: <https://www.pewresearch.org/> (accessed 3 June 2021).
- Silverman C (2016) This analysis shows how viral fake election news stories outperformed real news on Facebook. Available at: <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook#.pnEqX1D5R> (accessed 3 August 2021).
- Smarandache IG, Maricutoiu LP, Ilie MD, et al. (2021) Students' approach to learning: Evidence regarding the importance of the interest-to-effort ratio. *Higher Education Research & Development*. Epub ahead of print 10 January 2021. DOI: 10.1080/07294360.2020.1865283.
- Tapola A, Veermans M and Niemivirta M (2013) Predictors and outcomes of situational interest during a science learning task. *Instructional Science* 41(6): 1047–1064.
- Taylor A and Dalal HA (2017) Gender and information literacy: Evaluation of gender differences in a student survey of information sources. *College & Research Libraries* 78(1): 90–113.
- Tepe R and Tepe C (2015) Development and psychometric evaluation of an information literacy self-efficacy survey and an information literacy knowledge test. *Journal of Chiropractic Education* 29(1): 11–15.
- Vickery S and Cooper H (2003) Confidence or competence? Auditing information literacy skills of biology undergraduate students. In: *EDUCAUSE in Australasia 03*, Adelaide, SA, 6–9 May.
- Walsh A (2009) Information literacy assessment: Where do we start? *Journal of Librarianship and Information Science* 41(1): 19–28.

### Author biographies

Ellen Nierenberg is a PhD student in information literacy at UiT The Arctic University of Norway.

Tove I. Dahl is a professor in the Department of Psychology at UiT The Arctic University of Norway.

**Paper 3: Appendices**

## **Paper 3, Appendix 1: Estimated scores: Descriptive statistics and statistical analyses**

Statistics about students' estimated scores on the information literacy test are found in this supplement, rather than in the article, as we wished to focus on actual scores and estimation accuracy scores in the article. The table below shows values for actual test scores, estimated scores, and estimation accuracy scores, both before and after taking the IL test, for three higher education (HE) levels and two genders.

### **Estimated score comparisons**

A two-way ANOVA of estimated IL scores before the test by HE level and gender showed a main effect for HE level,  $F(2, 297) = 8.14, p < .001, \eta^2 = 0.05$ , but not for gender. Bonferroni post hoc tests showed that master's and PhD student estimations were significantly higher than the undergraduate estimations ( $p < .05$ ), but not significantly different from each other. A second two-way ANOVA of estimated scores after the test, by HE level and gender, showed a main effect for HE level,  $F(2, 741) = 81.15, p < .001, \eta^2 = 0.18$ , and gender,  $F(1, 741) = 10.90, p < .001, \eta^2 = 0.02$ . The HE level by gender interaction was statistically significant,  $F(2, 741) = 7.78, p < .001, \eta^2 = 0.02$ .

The Bonferroni post hoc test for estimations after the IL test showed that master's and PhD students estimated significantly higher scores than undergraduates,  $p < .001$ , and that PhD students estimated significantly higher scores than master's students,  $p < .05$ . Men estimated significantly higher scores than women in the undergraduate and master's groups,  $p < .001$ , but there was no significant gender difference in the PhD group.

Table: IL Test by HE Level and Gender: Descriptive Statistics for Test Scores, Estimated Scores and Estimation Accuracy Scores

Group (median)	Test score <sup>a</sup>			Estimated score						Estimation accuracy score <sup>c</sup>					
	<i>M</i>	<i>SD</i>	<i>n</i>	Before <sup>b</sup>			After			Before			After		
<i>M</i>				<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	
Undergrad. (13)	<b>12.71</b>	<b>3.12</b>	<b>330</b>	<b>14.81</b>	<b>4.38</b>	<b>70</b>	<b>11.95</b>	<b>5.00</b>	<b>329</b>	<b>1.19</b>	<b>5.27</b>	<b>70</b>	<b>-0.74</b>	<b>4.06</b>	<b>329</b>
Low <sup>d</sup>	9.72	2.11	137	13.90	4.59	21	10.20	4.35	137	4.57	5.28	21	0.48	4.83	137
High <sup>e</sup>	15.43	1.44	145	15.20	4.03	41	13.48	3.41	144	-0.76	4.27	41	-1.94	4.09	144
Master's (17)	<b>16.06</b>	<b>2.49</b>	<b>196</b>	<b>17.28</b>	<b>3.61</b>	<b>122</b>	<b>15.50</b>	<b>3.09</b>	<b>196</b>	<b>1.22</b>	<b>4.15</b>	<b>122</b>	<b>-0.56</b>	<b>3.59</b>	<b>196</b>
Low	13.99	1.79	96	17.02	4.17	64	14.89	3.43	96	2.75	4.60	64	0.90	3.70	96
High	18.65	0.74	63	17.95	2.84	40	16.49	2.32	63	-0.62	2.72	40	-2.16	2.95	63
PhD (17)	<b>16.89</b>	<b>2.13</b>	<b>234</b>	<b>16.28</b>	<b>3.85</b>	<b>116</b>	<b>16.43</b>	<b>3.21</b>	<b>234</b>	<b>-0.73</b>	<b>4.43</b>	<b>116</b>	<b>-0.46</b>	<b>3.46</b>	<b>234</b>
Low	14.54	1.52	80	16.27	3.81	37	15.55	3.65	80	1.86	4.40	37	1.01	4.18	80
High	18.85	0.90	93	16.50	3.55	52	17.23	2.62	93	-2.38	3.50	52	-1.62	2.52	93
Gender															
Male	<b>14.83</b>	<b>3.24</b>	<b>227</b>	<b>16.65</b>	<b>4.26</b>	<b>117</b>	<b>14.96</b>	<b>3.72</b>	<b>276</b>	<b>0.90</b>	<b>4.80</b>	<b>117</b>	<b>0.13</b>	<b>3.93</b>	<b>276</b>
Female	<b>14.80</b>	<b>3.31</b>	<b>471</b>	<b>16.08</b>	<b>3.80</b>	<b>186</b>	<b>13.79</b>	<b>4.12</b>	<b>471</b>	<b>0.22</b>	<b>4.51</b>	<b>186</b>	<b>-1.01</b>	<b>4.01</b>	<b>471</b>

<sup>a</sup> For all respondents; <sup>b</sup> Although all respondents estimated their scores *after* the IL test, a smaller subset of students estimated their scores *before* the test; <sup>c</sup> *Estimation accuracy score* is the estimated score minus the actual score of those who estimated; <sup>d</sup> Below median group; <sup>e</sup> Above median group

A two-way ANOVA for *before-test* estimated scores by HE level and performance-level group (low- or high-scorers) indicated a significant difference between HE levels,  $F(2, 254) = 10.44, p < .001, \eta^2 = 0.08$ , but not between low- and high-scoring groups. A Bonferroni post hoc test for HE level showed that differences between undergraduate and the graduate levels estimations were significant ( $p < .05$  for both), while the difference between master's and PhD students' estimated scores was not.

On the other hand, for *after-test* estimated scores by HE level and performance level group, there was a significant difference between HE levels,  $F(2, 612) = 90.95, p < .001, \eta^2 = 0.23$ , and performance level groups,  $F(1, 612) = 45.13, p < .001, \eta^2 = 0.07$ , as well as the interaction between HE level and performance level groups,  $F(2, 612) = 3.36, p < .05, \eta^2 = 0.01$ . A Bonferroni post hoc test for HE level showed the same trend as for before-test estimated scores, where the differences between undergraduate and the graduate levels estimations were significant ( $p < .001$  for both), while the difference between master's and PhD students' estimated scores was not. As seen in the table, high-scoring groups at all HE levels tended to estimate higher scores than low-scoring groups, both before and after the test.

In sum, while scores estimated *before* the test in the low- and high-scoring groups were similar at all HE levels, estimations made *after* the test were more in line with their actual performance – lower for the low-performing groups and higher for the high-performing groups.

## Paper 3, Appendix 2: Descriptive statistics

These data are for the following questions, posed to 134 undergraduate and PhD students, before and/or after the test:

1. 'I am interested in being or becoming an information literate person'  
(before and after the test; 1 = Not true for me at all, 6 = Very true for me)
2. 'I need more skills in information literacy'  
(before and after the test; 1 = Not true for me at all, 6 = Very true for me)
3. 'Knowing myself, I will make the effort to develop stronger information literacy skills' (after the test; 1 = not likely at all, 6 = highly likely)

**Table 1**

*Descriptive statistics for low- and high-performing<sup>a</sup> students' interest in being or becoming information-literate people<sup>b</sup>*

	Before IL test			After IL test		
	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>
Undergraduates						
Low performers	4.59	0.94	21	4.53	1.01	17
High performers	5.32	0.82	41	5.35	0.70	40
PhD students						
Low performers	5.60	0.68	20	5.25	0.97	20
High performers	5.28	1.30	32	5.55	0.63	29

<sup>a</sup> Low performers are those who scored below the median on the IL test, and high performers are those who scored above the median.

<sup>b</sup>Question 1 above

**Table 2**

*Descriptive statistics for questions about students' desire to be/become information literate, need for more IL skills, and likeliness to pursue this felt need*

	Before IL test			After IL test		
	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>
<b>I am interested in being or becoming an information literate person.</b>						
Undergraduates	5.11	0.93	70	5.14	0.87	64
PhD students	5.44	1.04	64	5.38	0.80	60
<b>I need more skills in information literacy.</b>						
Undergraduates	4.59	1.17	70	5.13	0.86	64
PhD students	4.31	1.17	64	4.55	1.06	60
<b>Knowing myself, I will make the effort to develop stronger information literacy skills.</b>						
Undergraduates	-	-	-	4.09	1.05	64
PhD students	-	-	-	4.70	1.03	60



**Paper 4**

## **Paper 4**

Submitted to *College & Research Libraries*, August 30, 2022

**Title:** Knowing, doing, and feeling: A three-year, mixed-methods study of undergraduates' information literacy development

**Authors:** Ellen Nierenberg (main author), Mariann Solberg, Torstein Låg, Tove Dahl

### **Abstract:**

This article reports results of a mixed-methods study following the development of undergraduates' information literacy (IL) over three years. IL knowledge and skills in this sample ( $n = 116$ ) increased with time, as did IL attitudes when measured by interest and IL's perceived usefulness and importance. Correlations among students' IL knowledge, skills, and attitudes also increased with time, implying a progressively stronger integration of the three. Complementary interviews with 13 students revealed that they became more interested in being information literate. Some experienced an identity change as a result of this development, indicating that transformative IL learning can occur.

## Introduction

Academic librarians began teaching critical thinking skills to students long before the *post-truth* era, in which “objective facts are less influential in shaping public opinion than appeals to emotion and personal belief”.<sup>1</sup> The term *post-truth* became common after the 2016 US presidential election and Brexit referendum in the UK, when misinformation spread by public figures became more prevalent. This major challenge in our information environment makes it difficult to distinguish between fact and fiction, and confirms the importance of information literacy (IL) skills such as the ability to critically evaluate information.

IL is defined in this study as encompassing “the knowledge, skills, and attitudes needed to be able to discover, evaluate, and use information sources effectively and appropriately in order to answer questions, solve problems, create knowledge, and learn”.<sup>2</sup> These competencies are composed of cognitive, behavioral, and affective elements. They are essential in education, the workplace, and daily life, for navigating the information landscape in the post-truth world, for critically evaluating information, and for becoming responsible citizens and reflective, lifelong learners. In higher education (HE), academic librarians and teaching faculty contribute to students’ acquisition of necessary information competencies through IL instruction and assessment. Two important frameworks for this instruction are the Association of College and Research Libraries’ [ACRL] former *Information literacy competency standards for higher education*<sup>3</sup> and current *Framework for information literacy for higher education*<sup>4</sup>. The transition from the ACRL standards to the framework signals a shift in IL teaching from a checklist approach to a more conceptual approach to IL.

In this study, mixed methods are employed to examine the development of IL knowledge, skills, and attitudes – referred to as *knowing*, *doing*, and *feeling* – in undergraduates over three years. Mixed methods research involves employing and integrating quantitative and qualitative methods to reach deeper understandings of research questions.<sup>5</sup> Mixing occurs both as each strand’s data, results, and inferences informs the development of the next strand, and in the integrated analyses of final findings. New questions were created in both surveys and interviews if the need arose after analyzing previously collected data, for example in asking interviewees to explain unexpected survey results.

## ***Literature review***

Important concepts in our exploration of IL development are grounded in interest theory and transformative learning theory. This section provides a brief overview of relevant literature.

### **IL development**

Studies over time, with or without specific interventions, often focus on one particular facet of IL, such as Rosman et al., who found an increase in students' information-seeking knowledge over three semesters.<sup>6</sup> Broader studies, encompassing several aspects of IL development over time, are more challenging to undertake and less prevalent in the IL literature. There is a need for more research to fill this gap and better demonstrate students' retention and longer-term IL learning. Especially valuable are studies that assess IL *knowing*, *doing*, and *feeling*, to explore the evolution of relationships among the three over time. Pinto and Fernández-Pascual do this in their quantitative study measuring changes in affective dimensions (attitudes and motivations) and IL knowledge and skills for a set of IL competencies. When measuring these components before and after a one-semester intervention, they found positive, significant relationships with some, but not all, of the IL competencies. Scores for IL knowledge and skills were lower than for affective dimensions, including self-efficacy, showing a miscalibration between students' actual and believed skills.<sup>7</sup> We found, however, no longitudinal research following IL *knowing*, *doing*, and *feeling* over a longer period of time.

### **Interest and meta-awareness**

IL instructors strive to teach in the best possible manner to help students become information literate individuals. To explore its effect on IL learning, the measurement of *feeling* in this study is based mainly on interest. Interest is conceptualized as both the psychological state of an individual engaged in certain content, and their motivation to continue that engagement. Interest influences learning by increasing both attention, expended effort, cognitive functioning, and the realization of goals.<sup>8</sup>

Our operationalization of interest is informed by Hidi and Renninger's four-phase model of interest development. Their model describes ways in which interest changes over time, and thereby has implications for both learning motivation and teaching methods.<sup>9</sup> The model describes four qualitatively distinct phases of interest, from triggered situational interest (the most fleeting and situation-dependent) to well-developed individual interest (the most stable and independently-pursued). Each phase has a unique psychological architecture described by

the relationship between four cognitive and affective variables – Situation dependence, Positive affect, Competence, and Meaningfulness.<sup>10</sup>

The four-phase model was first operationalized in a reliable and conceptually valid way by Dahl and Nierenberg, who tested it both with (a) self-chosen objects of interest, such as playing the bassoon, and (b) a specific interest, namely interest in being or becoming information literate.<sup>11</sup> We use the questionnaire developed for this second test to measure IL interest in the current study.<sup>12</sup>

### **Transformative learning**

Transformative learning (TL) theory is used as a lens to explore IL learning. TL was introduced in the 1970's to describe perspective transformations in adults that can occur after critical reflection of previous life experiences, assumptions, and beliefs.<sup>13</sup> Mezirow believed that TL is a type of paradigm shift that can induce more comprehensive behavioral changes than other kinds of learning.<sup>14</sup> Mezirow's theory has received criticism for being too cognitively oriented and lacking social and emotional elements of learning, and for not adequately capturing what it is that gets transformed in TL.<sup>15</sup> Illeris suggests that it is a person's *identity* that gets transformed, and proposes a new and simpler definition of TL, namely "all learning which implies changes in the identity of the learner"<sup>16</sup>. Illeris adds that TL "implies a qualitatively new formation [in a learner] ... something more than the acquisition of new knowledge and skills... [which may] include changes and transformations in the learner's general experience and behavior".<sup>17</sup> He specifies that TL involves cognitive, emotional, and social dimensions of learning, and encompasses only identity changes that are "of a certain importance in relation to the mental totality of the individual"<sup>18</sup>.

Personal identity, according to psychoanalyst Erik Erikson, consists of our perceptions of both ourselves and how others judge us; it forms mainly during youth and remains stable throughout life. Illeris and others, including sociologists Bauman and Giddens, maintain however that personal identity can change throughout life, thus supporting Illeris' supposition that identity is something that can be transformed.<sup>19</sup>

TL can be differentiated from other types of learning in several ways. TL refers specifically to learning in adults, after critical self-reflection, that leads to profound changes in both our

perspectives and behavior.<sup>20</sup> In addition, TL is learning that changes our identity<sup>21</sup>, and that is irreversible<sup>22</sup>.

In their literature review examining TL within the context of academic libraries, Hooper and Scharf explore how TL theory can inform librarians' IL teaching. They note that the ACRL's framework for IL has clear references to TL, including *threshold concepts*, defined as "core or foundational concepts that, once grasped by the learner, create new perspectives and ways of understanding... [and] produce transformation within the learner".<sup>23</sup> Other studies from the perspective of the IL educator include Chatterjee and Nichols Hess, who discuss how TL can inform IL instructional practice in HE. Kenney studied TL in public library users, none of whom were students. Hucks et al. explored how teacher education students showed evidence of TL when teaching IL in practice, thus combining the perspective of the student and the instructor.<sup>24</sup> Other than Hucks et al., we have not found IL studies that discuss TL from the student perspective.

Our research fills gaps in the IL literature regarding the use of mixed methods, and in studying the development of relationships between IL knowledge, practice, and attitudes over time. Additionally, we reestablish the construct of interest, an important motivator for learning into the IL literature.<sup>25</sup> Our exploration of *student* learning through the lens of TL theory also fills a gap, as previous research has focused on transformative experiences of those who *teach* IL.

### ***Research questions***

There are two main research questions in this study:

1. How do undergraduates' IL knowledge (*know*), skills (*do*), and interest in being or becoming information literate (*feel*) – and their interaction – change during the first three years of HE?
2. Do students' identities change to some extent as a result of becoming more information literate, indicating that transformative learning has occurred?

### ***The current study***

The development of IL in undergraduates over time was studied using mixed methods. During the first half of the study, quantitative methods were employed. It gradually became clear, however, that these methods alone were not sufficient to answer the research questions.

Qualitative methods were therefore employed during the second half of the study, enabling us to explore students' perceptions of their IL development more deeply.

This article is divided into 3 sections: first the quantitative and then the qualitative studies, including methods and results, and lastly the mixed methods section, where the quantitative and qualitative findings are integrated and discussed.

## Quantitative

### *Methods*

#### **Design of study**

This study measures development in IL *knowing*, *doing*, and *feeling* in psychology undergraduates at the world's northernmost university, UiT The Arctic University of Norway (UiT). Four measures, described in Materials, were employed to quantitatively examine students' IL development over their first three years. These tools measure their (1) knowledge of core facets of IL (*know*), (2) skills in evaluating sources (*do*), (3) skills in using sources (*do*), and (4) interest in being or becoming information literate (*feel*).

Using these tools, data were collected cross-sectionally during students' first and sixth semesters. For some of the sixth-semester respondents, data had also been collected three years previously, at the start of their first semester. Although these paired-sample first-semester data were not included in the *cross-sectional* study, where data from the same individual cannot be repeated, they were used in a complementary *longitudinal, paired-sample* study together with sixth-semester data taken from the cross-sectional study for these respondents. The cross-sectional study thereby measured the development of different students at the same point in their educational trajectory, while the longitudinal study followed the same students over time to examine in more detail what they *know*, *do*, and *feel* related to IL. In this exploratory research, we identified trends and validated results by comparing results of the cross-sectional and longitudinal studies.<sup>26</sup>

#### **Materials**

The four measures employed to assess IL *knowing*, *doing*, and *feeling* make up the *Tromsø Information Literacy Suite* (TROILS), and are freely available for others to use.<sup>27</sup> Additional details about the development and testing of the measures, including evidence of their reliability and validity, are described in Nierenberg et al. and Dahl and Nierenberg.<sup>28</sup>

*Knowing.* The Tromsø Information Literacy Test (TILT) is a 21-item, multiple-choice, psychometrically evaluated test designed to assess undergraduates' knowledge of three core, source-based facets of IL: finding, evaluating, and using sources.<sup>29</sup> There are seven items in each facet, each with four alternative answers, one of which is most correct. Maximum test score is 21 points.

*Doing.* Students' IL skills in practice are assessed, using mandatory assignments, for (a) source evaluation and (b) source use. These two measures arguably have inherent validity, as they are designed by teachers partly to assess these abilities.

(a) *Source evaluation* scores are based on mandatory assignments in the first and sixth semesters. Students find scholarly sources for their first-semester final papers and their sixth-semester bachelor's theses and describe why they consider each to be a reliable, scholarly source. In this research, three randomly chosen sources from these assignments are analyzed based on three criteria:

1. Quality: how scholarly the source is, on a scale of 0 (not scholarly) to 3 (scholarly).
2. Variety: number of unique source-evaluation criteria in the annotation, for example authority, relevancy, or accuracy. Each specific criterion is worth one point.
3. Frequency: total number of instances source-evaluation criteria appear in the annotation. Each criterion is worth one point.

The total score is the sum of the quality, variety, and frequency scores for the students' three sources. Scoring was performed by three raters in the first assignment, then averaged. Interrater reliability was found sufficient.<sup>30</sup> A qualified rater from the first-semester assignment scored the sixth-semester assignment, so all ratings are comparable.

(b) *Source use* scores, with a maximum of 5 points, are based on students' first-semester final papers and sixth-semester bachelor's theses. For both assignments, students were required to incorporate scholarly sources and cite them in APA style. The overall use of sources was assessed using five criteria, each worth 0-1 point:

1. Are scholarly sources used to support arguments?



2. Are sources cited in the text when necessary?
3. Are the in-text citations written in correct APA style?
4. Is the reference list written in correct APA style?
5. Are all in-text citations listed in the reference list, and vice versa?

The assessment of source use in the first semester was performed by three raters. Raters calibrated their scoring to ensure reliability, and interrater reliability was found satisfactory.<sup>31</sup> In the sixth semester, two raters assessed assignments and interrater reliability was again found sufficient.<sup>32</sup>

*Feeling* was defined in terms of interest in being or becoming an information literate person, and measured by the Tromsø Interest Questionnaire (TRIQ).<sup>33</sup> TRIQ consists of a self-assessed phase of interest and six theoretically-based subscales: General interest, Situation dependence, Positive affect, Competence level, Competence aspiration, and Meaningfulness. The subscales are derived from key variables from Hidi and Renninger's four-phase model of interest development.<sup>34</sup>

### **Participants**

All psychology students in this study have the same IL instruction, starting with an academic writing course with embedded IL in their first semester. When research began in fall 2019, students from two programs, a three-year bachelor's program ( $n = 75$ ) and a six-year professional studies program ( $n = 52$ ), were enrolled in this course. In their third year, all had instruction in literature search.

*Cross-sectional.* Data for the cross-sectional study were collected from cohorts that began in 2017-2019. For the 2019-cohort, data were analyzed from those students who participated in the survey (including TILT and TRIQ) in their first or sixth semester. For those who completed the survey in both semesters, only sixth semester data were included in the cross-sectional study, so no data from the same student were duplicated. The 2019-cohort had two mandatory assignments in their first and sixth semesters that were used to measure practical skills in source evaluation and source use (*do*). Bachelor's students in cohorts that began in 2017 and 2018 ( $n = 80$ ) had one comparable source-use assignment at the end of their sixth semester, namely their bachelor's theses. Since these cohorts had the same IL instruction and

assessment criteria as the 2019 cohort, their theses were included in the cross-sectional, source-use data.

*Paired-sample.* Paired data in the longitudinal study is from the 2019-cohort, who were followed over three years. Thirty-three of these students completed the survey (*know, feel*) in their first and sixth semesters. For the assignment-based measures (*do*), there is paired data only from students in the bachelor's program: source-evaluation data for 16 students and source-use data for 18 students.

### **Procedure**

TILT (*know*) and TRIQ (*feel*) data from the 2019-cohort were collected at the start of their first semester and end of their sixth semester.<sup>35</sup> Students were recruited via their Learning Management System (LMS) and email, and were offered prizes for participation.

With students' consent, source-evaluation and source-use data (*do*) from mandatory, written assignments in the first and sixth semesters were linked together with TILT and TRIQ scores. The use of sources was assessed in the first six pages of students' final exams in semester 1 and bachelor's theses in semester 6, where the nature of the writing is the same.

Data analyses were performed in IBM SPSS Statistics. For the cross-sectional study, independent sample t-tests were employed to compare first- and sixth-semester data from the four measures, as the assumptions for parametric tests were fulfilled in most cases. When data were not normally distributed, the non-parametric Mann-Whitney U test was used for comparison to check if the same trends were present. In these cases, for example sixth-semester TILT scores, the significance of the change between semesters was comparable in the parametric and non-parametric analyses.

For the longitudinal study, with its smaller sample size, the non-parametric Wilcoxon sign rank test with exact statistic enabled us to do the simplest, conservative analysis and still determine whether there were changes of note between the first and sixth semesters.

## Results

### Cross-sectional analyses

Semester 1 and 6 scores on the *know*, *do*, and *feel* measures in the cross-sectional study are compared in Table 1.

**Table 1: Independent-sample t-test results. Cross-sectional data for *know*, *do*, and *feel* scores in semesters 1 and 6**

Measure	Semester 1			Semester 6			<i>t</i>	<i>p</i> <sup>a</sup>	Cohen's <i>d</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>			
TILT	64	12.20	2.44	52	15.69	2.31	7.842	<.001	1.46
Source evaluation	78	9.65	2.94	34	11.45	2.44	3.137	.002	0.65
Source use	73	3.01	1.11	93	3.70	0.76	4.791	<.001	0.75
TRIQ									
- Interest phase	63	2.44	0.84	52	2.44	0.94	-0.013	.990	0.00
Subscales									
- General interest	63	4.41	0.92	52	4.15	0.86	-1.550	.124	0.29
- Situation dep.	62	2.82	1.12	52	3.03	1.12	0.980	.329	0.18
- Positive affect	60	3.91	0.98	52	3.68	0.85	-1.290	.200	0.24
- Comp. level	62	2.78	0.99	52	3.56	0.97	4.181	<.001	0.79
- Comp. aspiration	62	4.74	0.98	52	4.03	1.17	-3.559	<.001	0.67
- Meaningfulness	60	4.52	0.89	52	5.12	0.95	3.407	<.001	0.65

<sup>a</sup> Two-tailed *p*

Independent sample t-tests showed significant growth over time for scores on TILT and the source-evaluation and source-use measures, as expected. Interest phase remained the same over time, while TRIQ subscales show varied results. Significant results include growth in Competence level and Meaningfulness, which could be expected, and a decline in Competence aspiration.

### Paired sample analyses

#### *Development over time*

Results of the Wilcoxon sign rank test for paired samples, to examine *know*, *do*, and *feel* over time in the longitudinal group, are found in Table 2.

**Table 2: Wilcoxon signed rank tests with exact statistic for paired-sample, semesters 1 and 6, *know*, *do*, and *feel* score comparisons**

Variable	N	Semester 1		Semester 6		Pos. rank	Tie	Neg. Rank	Z	Exact sig. <sup>a</sup>
		M	SD	M	SD					
TILT	33	13.30	2.36	15.76	2.39	24	6	3	-4.175	<.001
Source evaluation	16	10.10	3.45	11.58	1.80	10	0	6	-1.501	.140
Source use	18	3.43	1.20	3.86	0.70	11	1	6	-1.593	.116
TRIQ	33									
- Interest phase		2.48	0.76	2.30	0.88	7	15	11	-1.108	.318
Subscales										
- General interest		4.41	1.03	3.97	0.78	9	3	21	-2.307	.020
- Situation dependence		2.62	0.97	3.15	1.16	20	3	10	-2.061	.038
- Positive affect		3.96	0.91	3.55	0.81	9	3	21	-2.445	.013
- Competence level		3.08	0.84	3.46	0.89	17	4	12	-1.973	.048
- Competence aspiration		4.72	1.07	3.73	1.16	6	4	23	-3.786	<.001
- Meaningfulness		4.76	0.82	4.93	0.86	18	3	12	-0.866	.395

*Note:* Positive ranks indicate higher scores in semester 6 than semester 1 (growth), and negative ranks indicate lower scores in semester 6 than semester 1 (decline).

<sup>a</sup>2-tailed

As shown by mean scores and ranks, there were significant improvements for TILT (*know*) in the longitudinal group. Source-evaluation and source-use scores (*do*) show growth as well, but this growth is not statistically significant.<sup>36</sup> Finally, in terms of what students *feel*, results from TRIQ subscales indicate similar trends as in the cross-sectional group, with the same subscales increasing or decreasing over time. Meaningfulness increased slightly and students reported that their Competence level improved significantly. However, their General interest decreased significantly, along with their feelings of Positive affect and Competence aspirations. Situation dependence increased significantly.

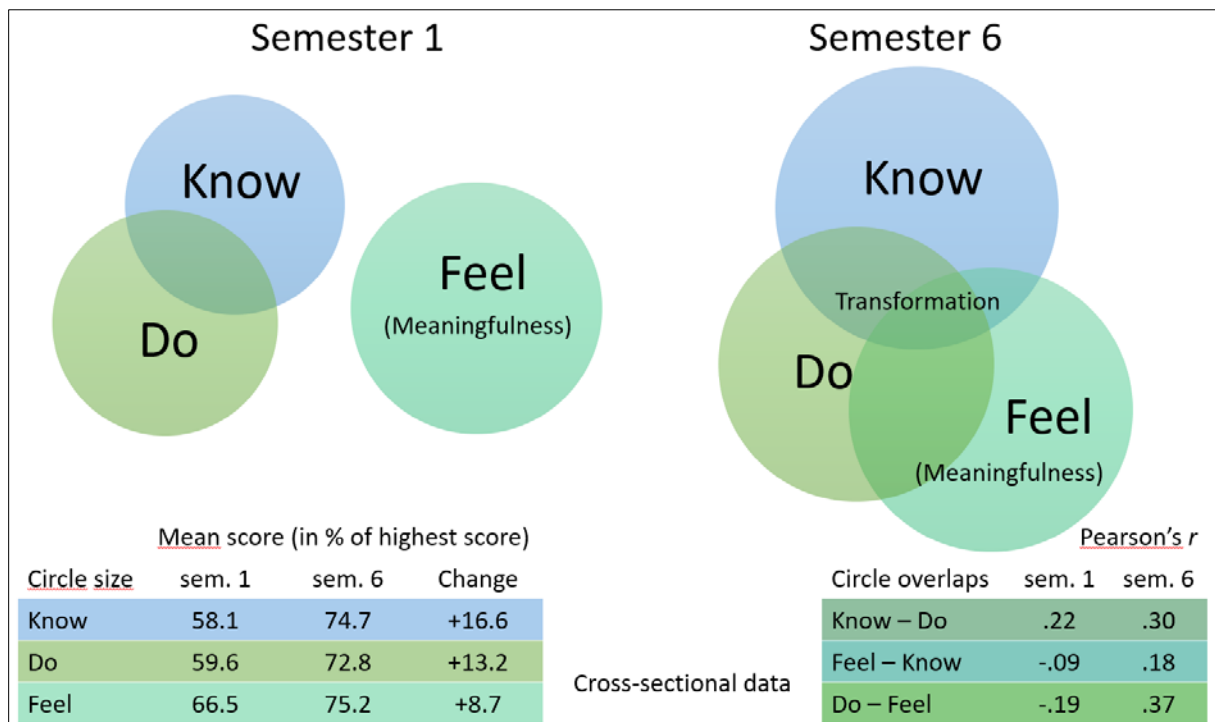
Data from cross-sectional and paired-sample groups are comparable, regarding both values and trends for each variable. The paired-sample results thereby corroborate findings from the cross-sectional sample.

### Comparisons between *knowing*, *doing*, and *feeling* over time

To examine how relationships between *knowing*, *doing*, and *feeling* developed between semesters 1 and 6 in the cross-sectional sample, scores on TILT, the two *do*-measures, and TRIQ's Meaningfulness subscale were analyzed. Since each TRIQ subscale measures a different aspect of interest, it gives no meaning to average them, as with the *do*-measures. The Meaningfulness subscale was chosen to represent overall interest because: (1) it has positive, significant correlations to students' interest phase in both semesters; (2) it increases most with interest phase, both in theory<sup>37</sup> and in practice<sup>38</sup>; (3) it contains items that are relevant to students' interest in being or becoming information literate, as well as to its usefulness and importance; (4) it is generally recognized in the field of educational psychology that both interest and the motivation to learn increase when activities are found meaningful.

Figure 1 is a qualitative illustration of IL growth, from semester 1 to 6, in the cross-sectional data. Circle size represents mean scores (bottom left table), the percent correct of the highest possible/achieved scores for *know* (TILT), *do* (average of source-evaluation and source-use), and *feel* (Meaningfulness subscale) measures. All three mean scores increased significantly over time. Overlapping areas between circles represent Pearson's  $r$  correlations (bottom right table), and illustrate areas where TL may occur (see Discussion). Correlations also increase, but these patterns of association are not particularly clear, and only the change in the *do-feel* correlation is statistically significant.<sup>39</sup> None of the six correlations are significantly different from zero, hence we cannot expect these patterns to hold for the population.

**Figure 1: Mean scores and correlations for knowing, doing, and feeling (Meaningfulness) in semesters 1 and 6, for cross-sectional data**



## Qualitative

### *Methods*

#### **Participants**

Qualitative data was collected in focus group interviews (FGI's) and individual interviews. Participants were sixth-semester, undergraduate psychology students from UiT, recruited via Canvas (learning management system) and email. Thirteen students participated in interviews – nine in FGI's and four in individual interviews.

During spring 2021, FGI's were conducted with students who began the psychology bachelor's program in 2018, who were then in the final semester of the three-year program, writing their bachelor's theses. There were three students in each of the three FGI's. Seven of the nine participants were women; the two men were in different groups. These students had the same curriculum and IL instruction as the cohort that began in 2019.

During spring 2022, individual interviews were conducted with four students who began studying psychology in 2019. Two were enrolled in the three-year bachelor's program and currently writing a bachelor's thesis, and two were in a six-year professional study program.<sup>40</sup>

## **Materials**

Interview guides for the FGI's and individual interviews differed slightly, but both included questions about students' perceptions of their IL development, the importance of IL for them, and their interest in being or becoming information literate people. Individual interviews focused additionally on TL.

A mixed-methods instrument, dubbed an "interest-o-meter", was also used during interviews.<sup>41</sup> This is an empty graph onto which participants plotted their levels of interest in being or becoming an information literate person during the past six semesters. The y-axis shows the intensity of the interest, and the x-axis is the timeline, in semesters. This graph, together with participants' simultaneous commentary and responses to other interest-related questions, provided us with a more detailed, mixed-methods account of their interest development.

## **Procedure**

Thirteen students volunteered and consented to participate in interviews. At the start of each interview, the procedure, the definition of IL used in the project, and information about the participants' privacy was presented. FGI's were conducted digitally by the main author in spring 2021, and all participants were provided the opportunity to answer each question. Individual interviews were conducted face-to-face in spring 2022. Interviews were semi-structured, providing the opportunity to pose follow-up questions and allowing students to introduce new topics. Interviews were audio recorded and transcribed verbatim, without emphasizing dialect or emotional expressions.

NVivo 12 was used for coding of interview data. The analysis was performed in several stages by the article's first two authors, using Braun and Clarke's<sup>42</sup> six-phase model for thematic analysis.<sup>43</sup> In the first stage of coding, an inductive approach was employed, creating codes to sort and describe the empirical data. In the next stage, codes were reorganized, splitting some and merging others. This stage was more abductive, informed by theories of TL and identity, and focused on the study's research questions. The subsequent creation of themes was based on these codes, with the intention of making meaningful patterns in the dataset that would contribute to answering research questions.

When confirming whether these themes were supported by a sufficient amount of data, however, we found that more data pertinent to TL was necessary. We therefore conducted follow-up individual interviews with three of the four participants from the first round, those who showed signs of TL. In this round, students were interviewed by the second author to see if new questions and a different interviewer's style could stimulate more detailed responses regarding any TL that may have occurred. The interview began by asking participants to plot their interest levels, per semester, on the interest-o-meter while commenting on their interest development and answering impromptu questions based on their visual representations. Subsequently, a question based on an unexpected survey finding was posed: "We have seen in the quantitative study that interest in being/becoming information literate decreased over time. Why do you think this might be the case?"

After this last round, we performed a thematic analysis based on codes from all interviews, identifying major themes in the data relevant to our research questions. Results of the thematic analysis and interest-o-meter are presented below.

### ***Results of thematic analysis***

Based on transcripts of focus group and individual interviews, 62 initial codes were created, with up to 30 references per code. After combining similar codes and semantically systematizing them into hierarchical categories, 28 main codes and 40 subcodes remained. Based on these, as well as the research questions and the theoretical framework, five major themes were found. These are described below, in alphabetical order, and illustrated with (translated) quotes from interviews. Names are changed to protect the anonymity of participants.

#### **A) Changes in identity**

Several students revealed that their perceptions of themselves (personal) and how others view them (social) – a common understanding of *identity* – had changed as a result of becoming more information literate. In the quote below, Amy's perception of herself has shifted; what she calls an *identity change*, and what Illeris would consider evidence of TL.<sup>44</sup>

Amy: I must have become [a more information literate person] without noticing it! When we talk about it now, absolutely, it has become a part of me, I have become very critical of what people say and where they get it from... It's the *situation* [university] and *practice*



[colloquia] that have played the biggest roles in my identity change... I'm very happy that I've become an information literate person.

While Amy describes her transition in an academic setting, Ruth and Heidi refer to daily life. Becoming critical of what people say and the sources they use represents a change toward a more critical way of thinking. This also depicts what Illeris would call a change in identity, indicative of TL.<sup>45</sup>

Ruth: After three years at university, I feel that I'm somewhat more competent than my parents... Becoming more information literate has made me a stronger person in a discussion context because I know how to express my opinions in an academic way.

Heidi: [I now] have less confidence in what I hear and see, and I have a greater need to verify... so I know what basis others have for what they say.

Other students also spoke about their identities changing to some extent as a result of becoming more information literate. Several felt more capable of arguing effectively, both in written work and in discussions with family and friends. They experienced that other people's perceptions of them had therefore changed, and that they were now being taken more seriously.

Critical self-reflection – questioning the validity of one's beliefs – is another indication of an identity change, and an important characteristic of TL. In the following quote, we see how Amy critically reflects on her previous behavior.

Amy: I now think twice before making assertions. I've always blurted out claims quickly... but now I'm better at waiting a little.

### **B) Development in IL Attitudes**

Participants felt that their attitudes about being information literate had developed during their undergraduate study. This evolution took place as a result of acquiring and using IL, including critical thinking skills, both in their studies and in daily life. Kari, who was writing her bachelor's thesis, expressed this change in attitude after illustrating her development as an information literate person on the interest-o-meter. Compared to her first year, she now cares

more about the quality of her work and about complying with the conventions of academic writing.

Kari: [I'm now] more involved and feel more ownership of what I write. I care much more about the product and will make sure that it's good quality.

Arthur reflects on his growth as a critical thinker, regarding both the consumption and creation of information.

Arthur: You learn a lot in just three years as a student, especially about becoming more critical of the information you receive.... [I now] think critically about things I post, for example in social media.

Interest in being or becoming information literate was often triggered by the need for IL skills, for example when writing papers. Ingrid's attitude about the importance of citing sources has evolved as she herself is becoming an information producer.

Ingrid: The more papers I write, the more I realize how important it is to that I [cite sources] correctly.

When asked which values and characteristics she generally associates with IL, Amy mentioned curiosity and an "open-minded critical approach." She felt that these characterized her as an information literate person.

### **C) Development in IL *Doing***

When asked how their abilities to find, evaluate, and use sources have developed, most students expressed substantial growth in all three areas. Source evaluation was where their greatest IL growth had occurred, followed by searching for sources. Those planning to continue their education expressed the need to learn even more about finding and evaluating sources because of their instrumental value; they still had use for these skills.

Learning to search more effectively for reliable sources represented an important breakthrough for several students. Ingrid was not alone in learning how to search for information other places than Google or Wikipedia.

Ingrid: In high school I thought that an information literate person searched in Wikipedia ... It was an *a-ha* experience for me when we learned about PsycINFO.<sup>46</sup>

Thor's abilities to find and evaluate information sources have developed greatly during HE.

Thor: I didn't know what it was to be critical to sources before I started college. I've really made a huge leap. Everything – from how I read, to where I choose to find information, to how I evaluate it – has changed, evolved.

Several participants told of great strides in learning why, when, and how to cite information sources. The quote below clearly shows Ingrid's development in reference technique in her written work.

Ingrid: Compared to my first year at the university, [citing sources] has become almost automatic. It's embarrassing to look at my reference lists then compared to now. APA style has become a good friend.

Plagiarism and its consequences were often mentioned as risks of *not* being able to cite sources correctly. Several participants, including Vera, were afraid of plagiarizing unintentionally, especially when writing in a language that is not their native tongue.

Vera: I hope I haven't plagiarized!... It's hard to know where to draw the line between plagiarism and paraphrasing... There's a gray zone there, especially in English, with my limited vocabulary. It's hard to find equally precise words as in the source.

#### **D) Development in IL *Knowing***

When inquiring about their IL knowledge development, all participants expressed considerable growth. Magnus began intentionally learning more about IL after realizing its usefulness and importance. He also mentions how becoming more knowledgeable can lead to a change in identity.

Magnus: When we had IL instruction, I suddenly realized how much I didn't know and have since been trying to learn more.... Obtaining more knowledge in general contributes to a change in one's identity.

Heidi and Vance exemplify additional risks of *not* being information literate and point to the difficulty of knowing which sources to trust.

Heidi: I'm knowledgeable enough that I don't drink bleach to avoid coronavirus when a president says so!

Vance: It's interesting to look at the consequences of *not* being information literate, how dangerous it can be to simply consume information without being critical, both in extremist and other political environments.

### **E) Interest in being or becoming information literate**

In TRIQ and interviews, instead of inquiring about students' interest *in IL*, we asked about their interest *in being or becoming information literate people*. The distinction between the two is small but significant, as the latter is relevant for one's identity. Although we expected student interest in IL to be fairly low, we thought that they might still recognize the value and importance of being or becoming information literate, and thereby have a potential interest in that aspect of their identity. This was the case for the students quoted below. For them, IL is mainly instrumental, something they need to write papers and pass their courses. They are not necessarily interested in *becoming* information literate, but recognize that by *being* information literate, they can achieve their goals.

Amy: I don't know if *interest* is a word I'd use to describe it, but I think it's *important* to be information literate.... [As a student,] it's important to *be*, but not necessarily to *become*, information literate.

Fred: I'm quite interested in *being* an information literate person... The main goal is to succeed as a student.

Eli's quote is an example of how interest in being information literate can be triggered by a certain situation where IL skills are required, how need can drive interest.

Eli: I thought much less about [being information literate] before I started studying... I felt that I didn't need to learn more about IL before we had [database search] instruction this semester. I benefited greatly by learning to search in databases.

### **Mixed methods**

#### ***Integrating findings***

Columns in Table 3 contain quantitative results from *know*, *do*, and *feel* (TRIQ subscales) measures in the cross-sectional study, and corresponding qualitative results and quotes from interviews.

**Table 3: Integrated results matrix: Undergraduates’ perspectives on their information literacy development**

Quantitative results	Qualitative results	Example quotes
TILT scores increased significantly.	When asked about the development of their IL knowledge, students described how they now know more about searching for, evaluating, and using information sources.	You learn a lot in just three years as a student. At least about becoming more critical of the information you find.
Scores on practical assignments involving the evaluation and use of sources increased significantly.	When asked about the development of their IL practice, students described being much better now at finding, evaluating, and using information.	I have developed several skills, especially how to cite correctly. And how to do good searches and evaluate sources, how good they are. I would say that is what has changed the most.
TRIQ scores decreased over time for General interest, which includes questions about interest in IL, and interest in being or becoming information literate.	When asked to speculate why survey results showed decreased General interest over time, students thought that IL may be considered by some to be most exciting when they first began learning about it, while it was still new.	The first year there were lots of people who were much more interested, it was probably a little more exciting, you were a little more ‘wide-eyed’. And now there are many who are bored. There are many who will not go on to a master’s and are tired of school.
TRIQ scores for Situation dependence increased slightly over time.	When asked the extent to which their interest in being or becoming information literate is dependent on the situation, students answered	If I further develop [my IL], then it is for use in work or study. It is mostly dependent on the situation that requires it.

	that yes, it is situation-dependent, both previously, currently, and in the future.	
TRIQ scores for Positive affect (pleasure, happiness, curiosity, engagement) decreased slightly over time.	When asked how feelings regarding their IL have changed over time, some students described initial growth, but then a gradual decrease in positive emotions.	I've been through the whole emotional register... [Now there is] less frustration and less joy, just a kind of tacit acceptance that this is how it is, this is how it must be done, and that it's important.
TRIQ scores increased significantly for Competence levels.	When asked how their IL abilities have developed with time, students spoke of significant improvements since beginning HE.	[As a new student] my IL competency was low, and I've since been trying to learn what I wasn't good at, like searching for information.  I am now more competent in finding sources and evaluating them. It's self-efficacy... Now I have a feeling of mastery.
TRIQ scores decreased significantly for Competence aspirations.	When asked to speculate about why Competence aspirations decreased in survey results, students described how the more they knew about IL, the less ambition they had to learn more.	[In the first year] there was potential for improvement, so I became more interested, because I wanted to reach a higher level.  [As a student] it is important to <i>be</i> , but not necessarily to <i>become</i> , information literate.

<p>TRIQ scores increased significantly for Meaningfulness.</p>	<p>When asked about how the meaningfulness of IL has changed for them, several students mentioned how it has become more meaningful regarding their present and future education and in everyday life, often referring to its instrumental value.</p>	<p><i>(In education)</i> I would describe my interest as very high now, especially since I'm working on my bachelor's thesis. There is a need for competence, and everything I've learned so far is being tested in this task. So I feel that [IL] is highly relevant and that I am interested in it.</p> <p>I now have an understanding of why it's so important to be information literate.</p> <p><i>(In everyday life)</i> Regarding avalanche knowledge, I have become very interested in this after coming to Tromsø and I feel IL is extremely relevant for deciding which sources to use, what is reliable.</p> <p><i>(Could it be a matter of life and death?)</i> Yes, it could. It's so extreme.</p>
--	---	---

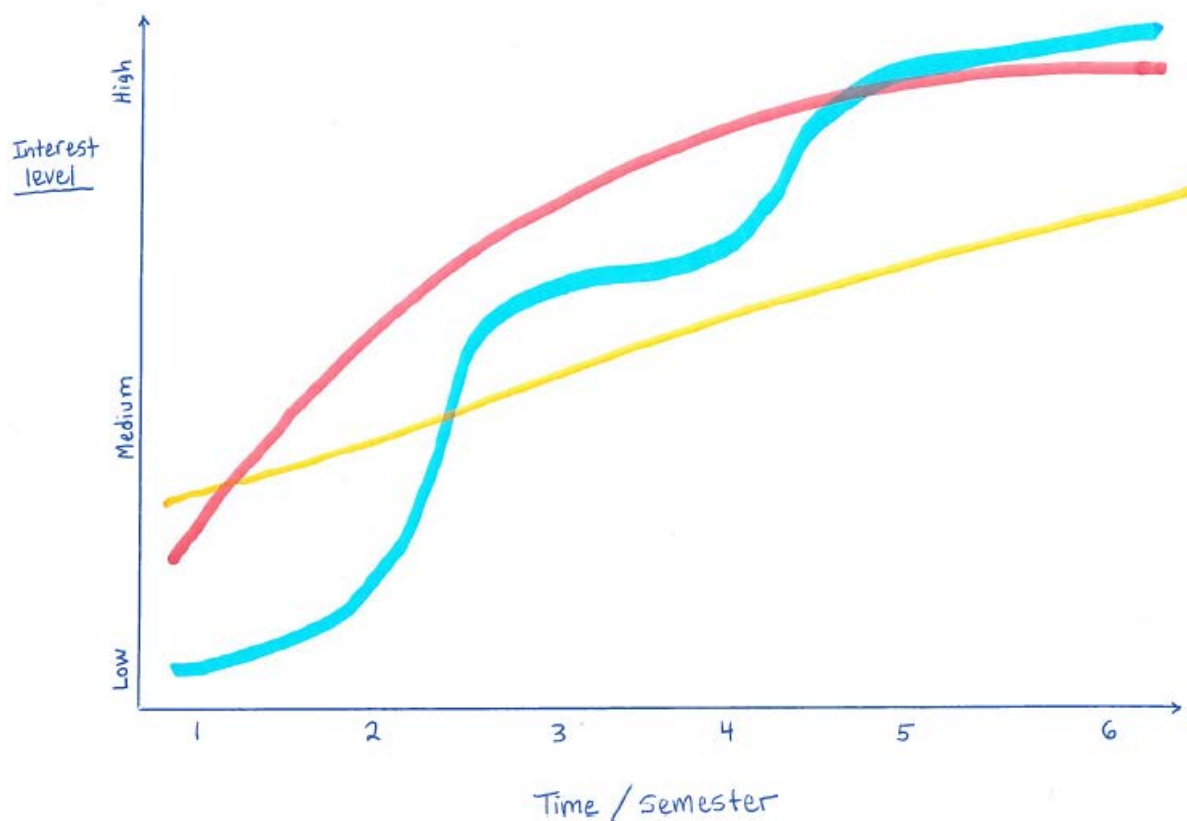
*Note.* In the right column, authors' words are in parentheses.



### ***Interest-o-meter***

Use of the interest-o-meter combined graphical, quantitative depictions of three students' interest in being or becoming information literate people over six semesters, with simultaneous oral accounts of this development. Figure 2 shows their drawings in the same graph. Kari (yellow) explained that her interest in the first semester was medium-low, but rose steadily to medium-high by the end of her sixth semester. Magnus (red) drew a gradually tapering upward curve depicting rapid initial growth and slower subsequent growth, with interest levels increasing from low-medium to high. He explained that he was more interested in being information literate than in IL in general. Amy's growth (blue), from low to high interest, was more cyclical, with jumps to higher plateaus each successive academic year. She commented that IL is interesting when important, but not as a subject in itself.

**Figure 2: Interest-o-meter showing three students' interest in being or becoming information literate people over six semesters**



### **Discussion**

#### ***Research questions***

The first research question in this study asks how *know*, *do*, and *feel*, and their interaction, change over the first three years of an undergraduate education. Regarding IL knowledge and

skills, we found – as expected – that both increased significantly with time in the quantitative data. Pinto and Fernández-Pascual, Rosman et al., and Scharf found similar results, although their studies followed students over only 1-3 semesters.<sup>47</sup> Analogous results were also found in interviews, where all informants experienced substantial growth in their IL knowledge and skills, articulating having grown mainly through *doing*. Assignments fueled their IL interest and the development of further knowledge and skills.

In addition to interviews, *feel* was assessed by the interest measure TRIQ. Tables 1 and 2 show some TRIQ subscales increasing and others decreasing with time. The quantitative increases in Competence level and Meaningfulness, which could be expected, were also expressed in interviews, where students felt that their IL knowledge and skill levels increased simultaneously with IL's meaningfulness for them. Students expressed interest in acquiring IL knowledge and skills when these were necessary for specific tasks; the need for skills fueled their interest.

Other subscale changes were more unexpected, such as declines in General interest (in IL and in being/becoming information literate) and Competence aspiration (desire to learn and do more with IL). This may be due to students' lesser need over time to acquire additional IL skills, beyond those already learned. Other unexpected subscale variations can be explained by students' changing interest phases (see Tables 1 and 2). For example, in the four-phase model of interest, lower interest phase signifies lower Positive affect and higher Situation dependence, which we observed.<sup>48</sup> Paradoxically, while interest phase declined or remained stable when measured quantitatively, interviewees expressed *increased* interest over time, also in interest-o-meters. This discrepancy may be due in part to interviewees being more motivated than others, as indicated both by their participation in interviews, their felt social expectations to communicate positive growth, or the fact that several intended to continue with graduate studies, and thereby had a further need for their IL skills.

Interactions between *knowing*, *doing*, and *feeling* over time are shown by correlations in Figure 1, as well as in interview material. When using the Meaningfulness subscale as a quantitative measure for *feel*, all correlations increased with time. The concurrent acquisition of IL knowledge, skills, and attitudes – and their interaction – together with results from the qualitative study, supports our assertion that TL regarding IL may have taken place, as discussed below.

As for the second research question, students' perceptions of themselves as information literate people have clearly evolved after three years in HE. This is evident both in TRIQ responses, where students reported substantial growth in their perceived IL abilities and in IL's meaningfulness for them, and in interview responses (see Table 3), where students spoke of feeling more competent and confident in their abilities to find, evaluate, and use sources, and in their ability to argue effectively. But is this learning *transformative*?

We believe that there is evidence of TL in this study, and base this claim on the defining characteristics of TL stated earlier in the article. TL refers specifically to learning:

- (a) *...in adults*. Third-year undergraduates are young adults with responsibility for their own learning.
- (b) *...after critical self-reflection*. Students told that, after three years of HE, they now question the validity of their previous beliefs and values. They have become less prone to believing and spreading false information and more aware that their previous methods of searching for information were ineffective. They better understand the importance of citing sources properly and have more integrity as knowledge creators.
- (c) *...that leads to profound changes in perspectives and behavior*. Several students told of behavioral changes in how they find, consume, and communicate information, and that they have become critical thinkers – basing their judgements and actions on reliable sources and deliberate, sound reasoning rather than emotional responses.
- (d) *...that leads to changes in identity*. In interviews, some students' words described identity changes – both how they perceive of themselves and their perceptions of how others judge them.<sup>49</sup> They feel more competent and confident in their IL abilities, use them in their education and daily life, and observe that others take them more seriously now that they base their arguments on reliable information.
- (e) *...that is irreversible*. When asked in interviews whether they could forget what they had learned about IL, students replied that the major tenets of IL, such as the importance of critically evaluating information sources, would always be remembered. (Future research would be required to determine whether this indeed is the case.)
- (f) *... implies a qualitative change in the learner, beyond the acquisition of new knowledge and skills*.<sup>50</sup> Several respondents expressed precisely this type of change, as exemplified by the selected quotes.

By employing mixed methods, where quantitative and qualitative findings are integrated and reinforce each other, the evidence for TL is strengthened. Quantitatively, TL is represented by the increase in correlations over time between *knowing*, *doing*, and *feeling* results (see circle overlaps in Figure 1). This indicates that they become more tightly integrated with time, an integration that can be interpreted within a framework of TL or *transformative integration*.<sup>51</sup> We imagine transformations taking place as this integrated whole of cognitive, behavioral, and affective dimensions gradually becomes part of one's identity. Qualitative results strengthen this supposition. Several students describe identity changes (i.e., transformations) as they became information literate people, including changes in their attitudes and behavior. Baartman and de Bruijn posit that this integration can involve a *transfer* of learning – applying what is learned in one task to a different task.<sup>52</sup> Informants indicated transfer when they spoke of broadly applying their IL abilities beyond academia. IL instructors may benefit from taking IL's transferability into consideration when teaching.

### ***Limitations***

A limitation of the quantitative study is the relatively small sample size for the source-evaluation measure (*do*) in the sixth semester. In the cross-sectional sample, only some students – those in the bachelor's program – had a source-evaluation assignment that semester. In the paired sample, few of the original students continued studying for three consecutive years, likely because of the pandemic.

Although qualitative research in general does not strive to be representative, it is worth mentioning, as a limitation, that interviewees were possibly more interested in IL than others in their cohort, given that they volunteered for interviews. Also, a limitation with interviews that require retrospective assessments of development is that respondents may not accurately recollect past experience. Had we interviewed students at separate points along their trajectories, instead of just their sixth semester, we may have obtained different results.

### ***Future research***

Ideas for future research include (1) comparing results from students who have IL instruction integrated in their studies with those who have one-shot instruction; (2) comparing those who have IL instruction in the first and second semester; (3) comparing results from IL measures with students' grades or completion rates to determine how IL levels are associated with

actual performance; and (4) studying growth with qualitative interviews along the entire student trajectory rather than just at the end.

### **Conclusion**

This article contributes to the body of IL research in several ways: (1) by following the development of undergraduates over three years; (2) by its use of mixed methods; (3) by assessing three aspects of IL development in parallel: knowledge, skills, and attitudes (as measured by students' interest in being or becoming information literate people); and (4) by looking for signs of transformative learning in students.

IL is a heterogeneous construct encompassing cognitive, behavioral, and affective elements, and research aimed at understanding IL development should therefore strive to touch upon all of these. Our mixed methods design, with its pragmatic, emergent nature, proved valuable in answering our complex research questions. The integration of quantitative and qualitative data enabled a more comprehensive description of students' IL development than either method on its own would have achieved.

For those of us who teach IL, this research shows that our efforts are worthwhile. Findings indicate firstly that students' IL knowledge, skills, and attitudes – when measured by the meaningfulness of being or becoming information literate – increase with time, as do the associations among them. Secondly, students become more critical to information and self-reflective, and may undergo a change in identity as a result of becoming more information literate. Such changes are indicative of transformative learning, a perspective transformation where individuals re-evaluate previous beliefs and assumptions to create new insight and meaning, a result that many educators strive for. Thirdly, relevant tasks in which IL skills are incorporated and assessed fuel students' interest and thereby their motivation to learn these IL skills. This is an important point to remember when designing IL instruction and assessment in the context of HE.

## Notes

<sup>1</sup> Oxford Dictionary, "Post-truth," accessed June 6, 2021,

<https://www.lexico.com/definition/post-truth>

<sup>2</sup> Ellen Nierenberg, Torstein Låg, and Tove I. Dahl, "Knowing and doing: The development of information literacy measures to assess knowledge and practice," *Journal of Information Literacy* 15, no. 2 (2021): 79, <https://doi.org/10.11645/15.2.2795>.

<sup>3</sup> "Information literacy competency standards for higher education," 2000, <http://hdl.handle.net/11213/7668>.

<sup>4</sup> "Framework for information literacy for higher education," 2015, <https://www.ala.org/acrl/standards/ilframework>.

<sup>5</sup> Abbas Tashakkori, Burke Johnson, and Charles Teddlie, *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*, 2nd ed. (Los Angeles, CA, USA: SAGE, 2020).

<sup>6</sup> Tom Rosman, Anne-Kathrin Mayer, and Günter Krampen, "A longitudinal study on information-seeking knowledge in psychology undergraduates: Exploring the role of information literacy instruction and working memory capacity," *Computers & Education* 96 (2016), <https://doi.org/10.1016/j.compedu.2016.02.011>.

<sup>7</sup> Maria Pinto and Rosaura Fernández-Pascual, "How a cycle of information literacy assessment and instruction stimulates attitudes and motivations of LIS students: A competency-based case study," *Journal of Librarianship and Information Science* 51, no. 2 (2019), <https://doi.org/10.1177/0961000617742447>.

<sup>8</sup> K. Ann Renninger and Suzanne E. Hidi, "Interest development and learning," in *The Cambridge Handbook of Motivation and Learning*, ed. K. Ann Renninger and Suzanne E. Hidi, Cambridge Handbooks in Psychology (Cambridge: Cambridge University Press, 2019); Judith M. Harackiewicz et al., "Predicting success in college: A Longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation," *Journal of Educational Psychology* 94, no. 3 (2002), <https://doi.org/10.1037/0022-0663.94.3.562>.

<sup>9</sup> Taehee Kim and Diane L. Schallert, "Mediating effects of teacher enthusiasm and peer enthusiasm on students' interest in the college classroom," *Contemporary Educational Psychology* 39, no. 2 (2014), <https://doi.org/10.1016/j.cedpsych.2014.03.002>; K. Ann

Renninger and Suzanne Hidi, *The power of interest for motivation and engagement* (New York: Routledge, 2016).

<sup>10</sup> Suzanne Hidi and K. Ann Renninger, "The four-phase model of interest development," *Educational Psychologist* 41, no. 2 (Spring 2006), [https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4).

<sup>11</sup> Tove I. Dahl and Ellen Nierenberg, "Here's the TRIQ: The Tromsø Interest Questionnaire based on the four-phase model of interest development," *Frontiers in Education* 6, no. 402 (2021), <https://doi.org/10.3389/feduc.2021.716543>.

<sup>12</sup> Available on <https://site.uit.no/troils>

<sup>13</sup> Jack Mezirow, "Perspective transformation," *Adult Education* 28, no. 2 (1978), <https://doi.org/10.1177/074171367802800202>.

<sup>14</sup> Jack Mezirow, *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning* (San Francisco, CA: Jossey-Bass, 1990).

<sup>15</sup> Knud Illeris, "Transformative learning and identity," *Journal of Transformative Education* 12, no. 2 (2014), <https://doi.org/10.1177/1541344614548423>; "Transformative learning as change and development of identity," in *Transformative learning meets bildung*, ed. Anna Laros, Thomas Fuhr, and Edward W. Taylor, International Issues in adult Education (Rotterdam: SensePublishers, 2017).

<sup>16</sup> Illeris, "Transformative learning and identity," 40.

<sup>17</sup> Illeris, "Transformative learning as change and development of identity," 179-80.

<sup>18</sup> Illeris, "Transformative learning and identity," 153.

<sup>19</sup> Erik H. Erikson, *Identity: Youth and crisis* (London: Faber & Faber, 1968); Illeris, "Transformative learning as change and development of identity."; Zygmunt Bauman, *Liquid modernity* (Cambridge: Polity Press, 2000); Anthony Giddens, *Modernity and self-identity: Self and society in the late modern age* (Cambridge: Polity Press, 1991).

<sup>20</sup> Mezirow, *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning*.

<sup>21</sup> Illeris, "Transformative learning and identity."; Illeris, "Transformative learning as change and development of identity."

<sup>22</sup> Chad D. Hoggan, "Transformative learning as a metatheory: Definition, criteria, and typology," *Adult Education Quarterly* 66, no. 1 (2016), <https://doi.org/10.1177/0741713615611216>; Edmund O'Sullivan, Amish Morrell, and Mary

Ann O'Connor, eds., *Expanding the boundaries of transformative learning: Essays on theory and praxis* (New York: Palgrave, 2002).

<sup>23</sup> Michaela D. Willi Hooper and Emily Scharf, "Connecting and reflecting: Transformative learning in academic libraries," *Journal of Transformative Education* 15, no. 1 (2017), <https://doi.org/10.1177/1541344616670033>; Association of College and Research Libraries, "Framework for information literacy for higher education," 9.

<sup>24</sup> Aneliia I. Chatterjee, "Using transformative learning theory to explore adjunct faculty's experiences with teaching information literacy: A case study of community colleges in New Jersey" (Doctor of Education Rowan University, 2022); Amanda Nichols Hess, "Academic librarians' teaching identities and work experiences: Exploring relationships to support perspective transformation in information literacy instruction," *Journal of Library Administration* 60, no. 4 (2020), <https://doi.org/10.1080/01930826.2020.1721939>; Brian Kenney, "The transformative library: A narrative inquiry into the outcomes of information use" (Ph.D., University of North Texas, 2008) (3376039); Darrell Hucks, Patrick Hickey, and Matthew Ragan, "Collaborations for transformative learning experiences: Technology integration and information literacy in teacher education," *International Journal of Cyber Ethics in Education* 4, no. 1 (2016), <https://doi.org/10.4018/IJCEE.2016010102>.

<sup>25</sup> Kim and Schallert, "Mediating effects of teacher enthusiasm and peer enthusiasm on students' interest in the college classroom."; Renninger and Hidi, *The power of interest for motivation and engagement*.

<sup>26</sup> All data are available at <https://doi.org/10.18710/SK0R1N>

<sup>27</sup> The tools are freely available on <https://site.uit.no/troils> and in UiT Open Research Data: <https://doi.org/doi:10.18710/L60VDI>

<sup>28</sup> Nierenberg, Låg, and Dahl, "Knowing and doing: The development of information literacy measures to assess knowledge and practice."; Dahl and Nierenberg, "Tromsø Interest Questionnaire."

<sup>29</sup> We do not imply that IL consists exclusively of these facets, only that they are crucial to the IL construct, as shown in the framework analysis of IL definitions, standards, and frameworks. See Nierenberg, Låg, and Dahl, "Knowing and doing: The development of information literacy measures to assess knowledge and practice."

<sup>30</sup> 0.89, see Nierenberg, Låg, and Dahl, "Knowing and doing: The development of information literacy measures to assess knowledge and practice," 88.



- <sup>31</sup> see Nierenberg, Låg, and Dahl, "Knowing and doing: The development of information literacy measures to assess knowledge and practice," 88.
- <sup>32</sup> Interrater reliability was .77 when measured with weighted Cohen's kappa.
- <sup>33</sup> TRIQ, and evidence of its reliability and validity, are found in Dahl and Nierenberg, "Tromsø Interest Questionnaire."
- <sup>34</sup> Hidi and Renninger, "The four-phase model of interest development."
- <sup>35</sup> TILT and TRIQ were included in a Qualtrics survey.
- <sup>36</sup> This is likely a result of the smaller sample size.
- <sup>37</sup> Hidi and Renninger, "The four-phase model of interest development."
- <sup>38</sup> see figures 1 and 4 in Dahl and Nierenberg, "Tromsø Interest Questionnaire."
- <sup>39</sup> Fisher's  $z = -1.962$ ,  $p = 0.049$
- <sup>40</sup> One male and one female from each of the two programs participated in individual interviews.
- <sup>41</sup> The interest-o-meter was modeled after Audun Hetland and Joar Vittersø, "The feelings of extreme risk: Exploring emotional quality and variability in skydiving and BASE jumping," *Journal of Sport Behavior* 35, no. 2 (2012). <https://hdl.handle.net/10037/2369>
- <sup>42</sup> Virginia Braun and Victoria Clarke, "Thematic analysis," ed. H. Cooper et al., vol. 2. *Research designs: Quantitative, qualitative, neuropsychological, and biological, APA handbooks in psychology: APA handbook of research methods in psychology* (Washington, DC: American Psychological Association, 2012).
- <sup>43</sup> The six phases in this model are: 1. Familiarizing yourself with your data, 2. Generating initial codes, 3. Searching for themes, 4. Reviewing themes, 5. Defining and naming themes, and 6. Producing the report.
- <sup>44</sup> Erikson, *Identity: Youth and crisis*; Illeris, "Transformative learning and identity," 40.
- <sup>45</sup> Harvey Siegel, "Educating reason: Critical thinking, informal logic, and the philosophy of education," *Informal Logic* 7, no. 2 (1985), <https://doi.org/10.22329/il.v7i2.2706>; Illeris, "Transformative learning and identity."
- <sup>46</sup> PsycINFO is a database for the discipline of psychology.
- <sup>47</sup> Pinto and Fernández-Pascual, "How a cycle of information literacy assessment and instruction stimulates attitudes and motivations of LIS students: A competency-based case study."; Rosman, Mayer, and Krampen, "A longitudinal study on information-seeking knowledge in psychology undergraduates: Exploring the role of information literacy

instruction and working memory capacity."; Davida Scharf, "An intervention and assessment to improve information literacy" (Ph.D. Rutgers University, 2013) (3597749).

<sup>48</sup> Hidi and Renninger, "The four-phase model of interest development."

<sup>49</sup> cf. Erikson, *Identity: Youth and crisis*.

<sup>50</sup> Illeris, "Transformative learning as change and development of identity."

<sup>51</sup> Liesbeth K. J. Baartman and Elly de Bruijn, "Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence," *Educational Research Review* 6, no. 2 (2011), <https://doi.org/10.1016/j.edurev.2011.03.001>.

<sup>52</sup> Baartman and de Bruijn, "Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence."

## References

Association of College and Research Libraries. "Framework for Information Literacy for Higher Education." 2015, <https://www.ala.org/acrl/standards/ilframework>.

———. "Information Literacy Competency Standards for Higher Education." 2000, <http://hdl.handle.net/11213/7668>.

Bauman, Zygmunt. *Liquid Modernity*. Cambridge: Polity Press, 2000.

Braun, Virginia, and Victoria Clarke. "Thematic Analysis." In *APA handbooks in psychology: APA handbook of research methods in psychology*, edited by H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf and K. J. Sher Washington, DC: American Psychological Association, 2012.

Baartman, Liesbeth K. J., and Elly de Bruijn. "Integrating Knowledge, Skills and Attitudes: Conceptualising Learning Processes Towards Vocational Competence." *Educational Research Review* 6, no. 2 (2011): 125-34. <https://doi.org/10.1016/j.edurev.2011.03.001>.

Chatterjee, Aneliia I. "Using Transformative Learning Theory to Explore Adjunct Faculty's Experiences with Teaching Information Literacy: A Case Study of Community Colleges in New Jersey." Doctor of Education, Rowan University, 2022.

- Dahl, Tove I., and Ellen Nierenberg. "Here's the Triq: The Tromsø Interest Questionnaire Based on the Four-Phase Model of Interest Development." *Frontiers in Education* 6, no. 402 (2021): 1-17. <https://doi.org/10.3389/educ.2021.716543>.
- Erikson, Erik H. *Identity: Youth and Crisis*. London: Faber & Faber, 1968.
- Giddens, Anthony. *Modernity and Self-Identity: Self and Society in the Late Modern Age*. Cambridge: Polity Press, 1991.
- Harackiewicz, Judith M., Kenneth E. Barron, John M. Tauer, and Andrew J. Elliot. "Predicting Success in College: A Longitudinal Study of Achievement Goals and Ability Measures as Predictors of Interest and Performance from Freshman Year through Graduation." *Journal of Educational Psychology* 94, no. 3 (2002): 562-75. <https://doi.org/10.1037/0022-0663.94.3.562>.
- Hetland, Audun, and Joar Vittersø. "The Feelings of Extreme Risk: Exploring Emotional Quality and Variability in Skydiving and Base Jumping." *Journal of Sport Behavior* 35, no. 2 (2012): 154-80. <https://hdl.handle.net/10037/2369>
- Hidi, Suzanne, and K. Ann Renninger. "The Four-Phase Model of Interest Development." *Educational Psychologist* 41, no. 2 (Spring 2006): 111-27. [https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4).
- Hoggan, Chad D. "Transformative Learning as a Metatheory: Definition, Criteria, and Typology." *Adult Education Quarterly* 66, no. 1 (2016): 57-75. <https://doi.org/10.1177/0741713615611216>.
- Hooper, Michaela D. Willi, and Emily Scharf. "Connecting and Reflecting: Transformative Learning in Academic Libraries." *Journal of Transformative Education* 15, no. 1 (2017): 79-94. <https://doi.org/10.1177/1541344616670033>.
- Hucks, Darrell, Patrick Hickey, and Matthew Ragan. "Collaborations for Transformative Learning Experiences: Technology Integration and Information Literacy in Teacher Education." *International Journal of Cyber Ethics in Education* 4, no. 1 (2016): 16-31. <https://doi.org/10.4018/IJCEE.2016010102>.
- Illeris, Knud. "Transformative Learning and Identity." *Journal of Transformative Education* 12, no. 2 (2014): 148-63. <https://doi.org/10.1177/1541344614548423>.
- . "Transformative Learning as Change and Development of Identity." In *Transformative Learning Meets Bildung*, edited by Anna Laros, Thomas Fuhr and

- Edward W. Taylor. *International Issues in Adult Education*, 179-90. Rotterdam: SensePublishers, 2017.
- Kenney, Brian. "The Transformative Library: A Narrative Inquiry into the Outcomes of Information Use." Ph.D., University of North Texas, 2008 (3376039).
- Kim, Taehee, and Diane L Schallert. "Mediating Effects of Teacher Enthusiasm and Peer Enthusiasm on Students' Interest in the College Classroom." *Contemporary Educational Psychology* 39, no. 2 (2014): 134-44.  
<https://doi.org/10.1016/j.cedpsych.2014.03.002>.
- Mezirow, Jack. *Fostering Critical Reflection in Adulthood: A Guide to Transformative and Emancipatory Learning*. San Francisco, CA: Jossey-Bass, 1990.
- . "Perspective Transformation." *Adult Education* 28, no. 2 (1978): 100-10.  
<https://doi.org/10.1177/074171367802800202>.
- Nichols Hess, Amanda. "Academic Librarians' Teaching Identities and Work Experiences: Exploring Relationships to Support Perspective Transformation in Information Literacy Instruction." *Journal of Library Administration* 60, no. 4 (2020): 331-53.  
<https://doi.org/10.1080/01930826.2020.1721939>.
- Nierenberg, Ellen, Torstein Låg, and Tove I. Dahl. "Knowing and Doing: The Development of Information Literacy Measures to Assess Knowledge and Practice." *Journal of Information Literacy* 15, no. 2 (2021): 78–123. <https://doi.org/10.11645/15.2.2795>.
- O'Sullivan, Edmund, Amish Morrell, and Mary Ann O'Connor, eds. *Expanding the Boundaries of Transformative Learning: Essays on Theory and Praxis*. New York: Palgrave, 2002.
- Pinto, Maria, and Rosaura Fernández-Pascual. "How a Cycle of Information Literacy Assessment and Instruction Stimulates Attitudes and Motivations of Lis Students: A Competency-Based Case Study." *Journal of Librarianship and Information Science* 51, no. 2 (2019): 370-86. <https://doi.org/10.1177/0961000617742447>.
- Renninger, K. Ann, and Suzanne Hidi. *The Power of Interest for Motivation and Engagement*. New York: Routledge, 2016.
- Renninger, K. Ann, and Suzanne E. Hidi. "Interest Development and Learning." In *The Cambridge Handbook of Motivation and Learning*, edited by K. Ann Renninger and Suzanne E. Hidi. Cambridge Handbooks in Psychology, 265-90. Cambridge: Cambridge University Press, 2019.

- Rosman, Tom, Anne-Kathrin Mayer, and Günter Krampen. "A Longitudinal Study on Information-Seeking Knowledge in Psychology Undergraduates: Exploring the Role of Information Literacy Instruction and Working Memory Capacity." *Computers & Education* 96 (2016): 94-108. <https://doi.org/10.1016/j.compedu.2016.02.011>.
- Scharf, Davida. "An Intervention and Assessment to Improve Information Literacy." Ph.D., Rutgers University, 2013 (3597749).
- Siegel, Harvey. "Educating Reason: Critical Thinking, Informal Logic, and the Philosophy of Education." *Informal Logic* 7, no. 2 (1985). <https://doi.org/10.22329/il.v7i2.2706>.
- Tashakkori, Abbas, Burke Johnson, and Charles Teddlie. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. 2nd ed. Los Angeles, CA, USA: SAGE, 2020.

## **Appendix 1: Framework analysis**

## Appendix 1: Framework analysis

Author, publication date	Framework/ Standard	Includes learning outcomes?	Description	Components
Carol Kuhlthau (1988)	Information Search Process	no	model for the information search process, shaped by cognitive, affective, and physical behavior	6 stages: <ol style="list-style-type: none"> <li>1. Task initiation</li> <li>2. Topic selection</li> <li>3. Pre-focus exploration</li> <li>4. Focus formulation</li> <li>5. Information collection</li> <li>6. Search closure</li> </ol>
Mike Eisenberg, Bob Berkowitz, Doug Johnson (1998/2000/2010)	The Big6 (including 'ICT Skills Curriculum Based on the Big6 Skills Approach to Information Problem-Solving')	yes	information and technology skills for student success in solving information problems (also used in primary and secondary school)	6 main stages: <ol style="list-style-type: none"> <li>1. Task definition</li> <li>2. Information seeking strategies</li> <li>3. Location and access</li> <li>4. Use of information</li> <li>5. Synthesis</li> <li>6. Evaluation</li> </ol>
Association of College and Research Libraries (ACRL; 2000)	Information Literacy Competency Standards for Higher Education	yes	5 standards, containing 22 performance indicators with corresponding outcomes	5 standards. <i>The information literate student...</i> <ol style="list-style-type: none"> <li>1. determines the nature and extent of the information needed</li> <li>2. accesses needed information effectively and efficiently</li> <li>3. evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system</li> <li>4. individually or as a member of a group, uses information effectively to accomplish a specific purpose</li> </ol>

				5. understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally
The Australian and New Zealand Institute for Information Literacy (ANZIIL; 2004)	Australian and New Zealand Information Literacy Framework	yes	framework with 6 standards, derived from ACRL (2000) <i>Information Literacy Competency Standards for Higher Education</i>	6 standards. <i>The information literate person...</i> 1. recognizes the need for information and determines the nature and extent of the information needed 2. finds needed information effectively and efficiently 3. critically evaluates information and the information seeking process 4. manages information collected or generated 5. applies prior and new information to construct new concepts or create new understandings 6. uses information with understanding and acknowledges cultural, ethical, economic, legal, and social issues surrounding the use of information
Christine Bruce, Sylvia Edwards, Mandy Lupton (2006)	Six Frames for Information Literacy Education	no	conceptual framework for interpreting the relationships between theory and practice in IL education	6 frames: 1. Content Frame 2. Competency Frame 3. Learning to Learn Frame 4. Personal Relevance Frame 5. The Social Impact Frame 6. Relational Frame



Emma Coonan, Jane Secker (2011)	A New Curriculum for Information Literacy (ANCIL)	yes	curriculum for IL with 10 thematic strands, each containing facets with learning outcomes, activities, and assessments	10 strands: 1. Transition from school to higher education 2. Becoming an independent learner 3. Developing academic literacies 4. Mapping and evaluating the information landscape 5. Resource discovery in your discipline 6. Managing information 7. Ethical dimension of information 8. Presenting and communication knowledge 9. Synthesizing information and creating new knowledge 10. Social dimensions of information
Society of College, National and University Libraries (SCONUL; 2011)	The SCONUL Seven Pillars of Information Literacy	yes	model defining core skills and competencies (ability) and attitudes and behaviors (understanding) for IL development	7 pillars: 1. Identify 2. Scope 3. Plan 4. Gather 5. Evaluate 6. Manage 7. Present
United Nations Educational, Scientific and Cultural Organization (UNESCO; 2013)	Global Media and Information Literacy Assessment Framework: Country Readiness and Competencies	yes	Set of tools to assess media and information literacy (MIL)	Two-tier approach: 1. MIL country readiness (MIL education, MIL policy, MIL supply, MIL access and use, civil society) 2. MIL competencies (rights, attitudes, values, knowledge, skills)

<p>Association of College and Research Libraries (ACRL; 2015)</p>	<p>Framework for Information Literacy for Higher Education</p>	<p>no</p>	<p>framework with 6 frames; replaces ACRL Standards (2000); shifts the focus of IL instruction from skills to process</p>	<p>6 frames (alphabetically):  a) Authority is Constructed  b) Contextual Information Creation as a Process  c) Information has Value  d) Research as Inquiry  e) Scholarship as Conversation  f) Searching as Strategic Exploration</p>
---	--	-----------	---	--

**Appendix 2: Course Certificate**

## Appendix 2: Course Certificate

Health Research Act – Course Certificate from REK (Regional Committees for Medical and Health Research Ethics)

 helsebiblioteket.no

**E-læringskurs i**

Helseforskningsloven

**er gjennomført og bestått**

Ellen Nierenberg

14 April 2019



**Appendix 3: NSD notification form**

## Appendix 3: NSD notification form

**Notification form for personal data from The Norwegian Centre for Research Data (NSD). This form includes information about the research project, the individual samples, types of (personal) data collected, consent forms, data management, and data security.**

Meldeskjema 102240

Sist oppdatert 25.01.2021

Hvilke personopplysninger skal du behandle?

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidentifikator
- Lydopptak av personer

Type opplysninger

Skal du behandle særlige kategorier personopplysninger eller personopplysninger om straffedommer eller lovovertridelser?

Nei

### Prosjektinformasjon

Prosjektittel: Din informasjonskompetanse i høyere utdanning

Prosjektbeskrivelse:

I dette PhD-prosjekt skal jeg måle bachelor-studenters informasjonskompetanse og hvordan den utvikler seg over tid. Det første målet er å utvikle verktøy som kan måle både studentenes kunnskaper, ferdigheter og interesse når det gjelder informasjonskompetanse. En elektronisk survey skal måle kunnskaper og interesse. For å måle ferdigheter, skal jeg analysere to skriftlige oppgaver i forhold til kildekritikk (kritisk vurdering av kilder) og kildebruk (bruk av henvisninger og referanser). Det andre målet er å sammenligne resultater fra to utvalg: første-års psykologistudenter, som får integrert informasjonskompetanseundervisning, og første-års studenter fra andre emner hvor den ikke er integrert. Det tredje målet er å følge to kull psykologistudenter over to og tre år for å se på deres utvikling over tid. Studenter får dermed surveyen tilsendt via e-post tre eller fire ganger. Analyser av studentenes skriftlige arbeider blir gjort i deres første og sjette semestre.

Begrunn behovet for å behandle personopplysningene

For å analysere studentenes utvikling over tid må data samles i flere omganger. Studentene vil få en survey tilsendt via e-post 4 ganger i løpet av deres 3-årige bachelorprogram. Det blir derfor nødvendig å beholde e-post adressene til respondentene i 3 år. For å få tilgang til studentenes skriftlige innleveringer, må jeg ha registerdata fra UiTs studentsystem (FS) og læringsplattform (Canvas), inkludert identitet (navn og studentnummer), e-post og kandidatnummer.

Ekstern finansiering  
Type prosjekt  
Forskerprosjekt  
Behandlingsansvar

---

### **Behandlingsansvarlig institusjon**

UIT – Norges Arktiske Universitet / Det helsevitenskapelige fakultet / Institutt for psykologi  
Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)  
Ellen Nierenberg, ellen.nierenberg@uit.no, tlf: 99435850  
Skal behandlingsansvaret deles med andre institusjoner (felles behandlingsansvarlige)?  
Nei

### **Utvalg 1**

---

Beskriv utvalget

Studenter som begynner på emnet 'Tenkning, læring og skriving i høyere utdanning' (PSY-0700) ved UiT Norges arktiske universitet i 2019, 2020 eller 2021.

Rekruttering eller trekking av utvalget

Alle studenter i utvalg 1 inviteres til å delta. Invitasjon fremsettes muntlig på et informasjonsmøte og så sendes via e-post og ved nyhetsmelding i Canvas (UiTs læringsplattform).

Alder

18 - 100

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 1

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidifikator
- Lydopptak av personer

Hvordan samler du inn data fra utvalg 1?

Elektronisk spørreskjema

Vedlegg

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 1

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Informasjonsskriv

## Utvalg 2

---

Beskriv utvalget

Respondenter fra utvalg 1, årskull 2019, som tar emnet 'Bacheloroppgave' (PSY-2002) i sitt sjette semester.

Rekruttering eller trekking av utvalget

Alle studenter i utvalg 2 inviteres til å delta. Invitasjon fremsettes muntlig på et informasjonsmøte, og så sendes via e-post og ved nyhetsmelding i Canvas (UiTs læringsplattform).

Alder

18 - 100

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 2

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidentifikator
- Lydopptak av personer

Hvordan samler du inn data fra utvalg 2?

Elektronisk spørreskjema

Vedlegg

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 2

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Informasjonsskriv

## Utvalg 3

---

Beskriv utvalget

Første-års UiT studenter (årskull 2020) i emnene: - 'Examen Facultatum ved HSL-fakultetet' (EXF-0712 og EXF-0714) - 'Felleskurs Helsefakultetet' (HEL-0700) - 'Examen Facultatum, idrett' (IDR-0110) - 'Examen Paedagogicum' (PED-0700)

Rekruttering eller trekking av utvalget



Alle studenter i utvalg 3 og utvalg 1 (årskull 2020) inviteres til å delta. Invitasjon fremsettes muntlig på et informasjonsmøte, og så sendes via e-post og ved nyhetsmelding i Canvas (UiTs læringsplattform).

Alder

18 - 100

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 3

- E-postadresse, IP-adresse eller annen nettidentifikator
- Lydopptak av personer

Hvordan samler du inn data fra utvalg 3?

Elektronisk spørreskjema

Vedlegg

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 3

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Informasjonsskriv

#### **Utvalg 4**

---

Beskriv utvalget

Frivillige studenter fra emnet PSY-2002 ved UiT, både vår 2021 og vår 2022. Disse studentene er i det siste semesteret av sin Bachelor i psykologi. Jeg vil ha fokusintervjuer med 16 av dem (4x4) fra hvert årskull.

Rekruttering eller trekking av utvalget

Alle studentene i PSY-2002 vil få et forespørsler i Canvas om å delta på fokusintervju. De blir tilsendt et infoskriv om personvern.

Alder

19 - 99

Inngår det voksne (18 år +) i utvalget som ikke kan samtykke selv?

Nei

Personopplysninger for utvalg 4

- Navn (også ved signatur/samtykke)
- E-postadresse, IP-adresse eller annen nettidentifikator
- Lydopptak av personer

Hvordan samler du inn data fra utvalg 4?

Gruppeintervju

Vedlegg

Grunnlag for å behandle alminnelige kategorier av personopplysninger

Samtykke (art. 6 nr. 1 bokstav a)

Informasjon for utvalg 4

Informerer du utvalget om behandlingen av opplysningene?

Ja

Hvordan?

Skriftlig informasjon (papir eller elektronisk)

Informasjonsskriv

Tredjepersoner

---

Skal du behandle personopplysninger om tredjepersoner?

Nei

Dokumentasjon

---

Hvordan dokumenteres samtykkene?

- Elektronisk (e-post, e-skjema, digital signatur)

Hvordan kan samtykket trekkes tilbake?

For å trekke tilbake samtykket må studentene henvende seg til meg via e-post slik at jeg kan fjerne dem fra utvalget. I informasjonsskrivet ved hver utsending står det at de kan trekke seg når de vil, og de får min e-post adresse.

Hvordan kan de registrerte få innsyn, rettet eller slettet opplysninger om seg selv?

De får innsyn ved henvendelse til meg. Jeg kan endre eller slette opplysningene.

Totalt antall registrerte i prosjektet

100-999

Tillatelser

---

Skal du innhente følgende godkjenninger eller tillatelser for prosjektet?

Behandling

---

Hvor behandles opplysningene?

- Maskinvare tilhørende behandlingsansvarlig institusjon
- Mobile enheter tilhørende behandlingsansvarlig institusjon
- Ekstern tjeneste eller nettverk (databehandler)

Hvem behandler/har tilgang til opplysningene?

- Prosjektansvarlig

- Interne medarbeidere
- Student (studentprosjekt)
- Databehandler

Hvilken databehandler har tilgang til opplysningene?

Qualtrics

Tilgjengeliggjøres opplysningene utenfor EU/EØS til en tredjestat eller internasjonal organisasjon?

Nei

Sikkerhet

---

Oppbevares personopplysningene atskilt fra øvrige data (kodenøkkel)?

Ja

Hvilke tekniske og fysiske tiltak sikrer personopplysningene?

- Opplysningene anonymiseres
- Endringslogg
- Flerfaktorautentisering
- Adgangsbegrensning

Varighet

---

Prosjektperiode

01.03.2019 - 31.10.2022

Skal data med personopplysninger oppbevares utover prosjektperioden?

Nei, data vil bli oppbevart uten personopplysninger (anonymisering)

Hvilke anonymiseringstiltak vil bli foretatt?

- Lyd- eller bildeopptak slettes
- Koblingsnøkkelen slettes
- Personidentifiserbare opplysninger fjernes, omskrives eller grovkategoriseres

Vil de registrerte kunne identifiseres (direkte eller indirekte) i oppgave/avhandling/øvrige publikasjoner fra prosjektet?

Nei

Tilleggsopplysninger

---

De vedlagte spørreskjemaer er pilotundersøkelser. Resultater fra disse skal brukes i utarbeidelsen av de endelige spørreskjemaer (før august 2019). Det formelle forespørsel om samtykke fra respondentene finnes i de vedlagte informasjonsskrivene for hvert utvalg. Det som står som introduksjon i de vedlagte spørreskjemaene er en kortere versjon av informasjonsskrivet, og regnes ikke som det formelle forespørsel om samtykke.

**Appendix 4: Consent Form**

## Appendix 4: Consent Form

### Consent form

I followed students in the fall 2019 introductory psychology course for three years for my research. The main sections of the consent form below (in Norwegian) are: Background and aims of the study; What participation in the study entails; Benefits and risks; What happens to collected information; Voluntary participation; and Consent to participate in the study.

In this consent form I ask students' consent to follow their progress and to keep their e-mail addresses during this time period. I also inform them how I will protect their privacy. Consent forms for other student groups are similar. All consent forms were assessed by NSD, The Norwegian Centre for Research Data.

### *Forespørsel om deltakelse i forskningsprosjektet: Din informasjonskompetanse i høyere utdanning*

#### *Bakgrunn og formål* (Background and aims of the study)

Informasjonskompetanse (IK) er de ferdigheter, kunnskaper og holdninger vi trenger for å hensiktsmessig kunne bruke informasjonskilder til å svare på spørsmål, løse problemer og til å lære. IK-ferdigheter inkluderer å kunne finne og kritisk vurdere informasjonskilder, og å kunne bruke dem på korrekt måte i skriftlige arbeid. Men hvordan sørger vi for at studenter i høyere utdanning oppnår god IK? For å nå dette målet må vi ha god, forskningsbasert kunnskap om hva som påvirker utviklingen av IK hos studenter i høyere utdanning.

For å bidra til dette gjennomfører vi en større undersøkelse blant studenter ved UiT, som en del av min doktorgradsstudie. Vi ønsker å sammenholde data om både kunnskap, ferdigheter og interesse fra forskjellige utvalg med litt ulike IK-undervisningstilbud. I den forbindelse ønsker vi å spørre deg som er oppmeldt til 'Tenking, læring og skriving i høyere utdanning' (PSY-0700) om samtykke til å få tilgang til de arbeidskravene og eksamensbesvarelsene du leverer i Canvas. Samtykket gjelder både for PSY-0700 og for 'Bacheloroppgave' (PSY-2002) i det 3. året. Dette innebærer å koble ditt kandidatnummer på eksamen til ditt navn og ditt studentnummer i Canvas. I tillegg ønsker vi å følge din fremgang over tid ved hjelp av en kort survey som distribueres hvert år av studiet. I den forbindelsen spør vi derfor om lov til å beholde også din e-post adresse over tre år. Denne skal kun brukes til å sende deg surveyen.

#### *Hva innebærer deltakelse i studien?* (What does participation in the study entail?)

Dersom du samtykker til det, henter vi informasjonen vi trenger fra UiTs registre. I tillegg fyller du ut den korte surveyen når vi forespør deg om det.

#### *Fordeler og risikoer* (Benefits and risks)

Det er ingen risiko ved å delta, fordi alle dine personopplysninger er konfidensielle og lagret sikkert, og blir slettet når dataanalysen er ferdig. Fordelen er at du bidrar til viktig forskning om informasjonskompetanseundervisning, som vil gagne framtidige studenter.

*Hva skjer med informasjonen om deg? (What happens to the information about you?)*

Alle personopplysninger vil bli behandlet konfidensielt og lagret sikkert. Kun jeg og mine tre PhD-veiledere, Professor Tove I. Dahl, Førstebibliotekar Torstein Låg og Professor Mariann Solberg, vil ha tilgang til de innhentede opplysningene, som påføres en tallkode slik at de lagres i aidentifisert form. Koblingsnøkkelen lagres passordbeskyttet og adskilt fra øvrige opplysninger om deg. Data fra dette semesteret blir sammenstilt og analysert først etter at alle arbeidskrav og eksamen er gjennomført.

Resultatene fra prosjektet vil bli publisert i min doktoravhandling og i internasjonale tidsskrift, men på en slik måte at ingen enkeltpersoner kan kjennes igjen.

Prosjektet skal etter planen avsluttes 30.06.22. Da vil koblingsnøkkelen slettes, og data som lagres utover dette tidspunktet vil være anonymisert.

*Frivillig deltakelse (Voluntary participation)*

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert. Hvis du velger å ikke gi ditt samtykke, får det ingen konsekvenser for dine studier eller for ditt forhold til Institutt for psykologi eller UiT.

Dersom du har spørsmål til studien, ta kontakt med Ellen Nierenberg, e-post: ellen.nierenberg@uit.no tlf. 776 44 133.

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

*Samtykke til deltakelse i studien (Consent to participate in the study)*

Du samtykker ved å trykke på knappen merket 'Neste' nederst til høyre, deretter på radioknappen merket 'Ja', og til slutt på knappen 'Send inn svar'. Dersom du ikke vil delta, lar du være å svare. Lenke til surveyen finner du ved å trykke 'Neste' igjen, etter at du har levert samtykke.

(Spørsmålstekst i Canvas: Jeg har mottatt informasjon om studien, og samtykker til at opplysninger om meg kan innhentes fra Institutt for psykologi, UiT Norges arktiske universitet.)

# Appendix 5: Data Management Plan

# Appendix 5: Data Management Plan

Submitted to The Norwegian Centre for Research Data (NSD) on March 3, 2019, and last changed April 12, 2021. Translated from Norwegian.

## Information literacy in higher education

In this PhD project, I will measure undergraduates' information literacy (IL) and how it develops over time. The first aim is to develop tools that can measure both the students' knowledge, skills, and interest in IL. An electronic survey will measure knowledge and interest. To measure skills, I will analyze two written tasks for skills in source evaluation (critical evaluation of sources) and source use (in-text citations and references). The second aim is to compare results from two samples: first-year psychology students, who receive integrated IL instruction, and first-year students from other subjects where it is not integrated. The third aim is to follow two cohorts of psychology students over two and three years to look at their development over time. Students are thus sent the survey via e-mail three or four times. Analyses of the students' written work are done in their first and sixth semesters.

**Discipline:** Social sciences

**Affiliation:** UiT The Arctic University of Norway / Faculty of Health Sciences / Department of Psychology

## Description of data

### Project goals

The project's aim is three-fold: 1) to develop and validate instruments to measure IL, 2) compare levels of IL among students with different IL instruction, 3) measure the development of and interaction between interest and IL over time. Three instruments will be developed: a survey, an annotated bibliography, and a rubric. The survey has two measures - one regarding IL the other regarding interest. The project will collect and/or generate new data.

### Value of project

Research data from this project will be useful for (1) anyone interested in measuring IL in higher education, (2) those who study and measure interest, and (3) those who want to further develop measurement instruments. Anonymized data will be freely available, and researchers who want to build on and/or further develop the instruments will find them – and research data – on the project's open website.

### Keywords

- IL
- Interest



- Measuring instruments
- Higher education
- Source evaluation
- Source use
- Annotated bibliography
- Rubric
- Survey
- Library and information science
- Psychology
- Students

### **Language**

- English
- Norwegian

### **Technical information**

#### **Types of data**

- Data sets
- Text

Most will be tabular data (numerical or categorical). The raw data from the annotated bibliography will also include some text (in Norwegian).

#### **Estimated data size**

1 GB – 100 GB

#### **Method of data collection/generation**

- Qualtrics survey with link from LMS (Canvas) or e-mail
- annotated bibliography assignment, submitted in Canvas
- written assignment, rubric score on selected sections
- Self-administered questionnaire
- Content coding
- Self-administering text or log

#### **Format**

Data analyses will be performed mainly in R and SPSS. We intent to make data available in open and persistent (archival-worthy) file formats:

- tab-separated Unicode text (.txt) for spreadsheets
- PDF/A (.pdf) for text, figures, diagrams
- R-script and/or SPSS syntax for the data analyzes will also be made available

## **Ethical and legal considerations**

Personal data will be collected or generated. Data will be collected that may make people identifiable or that contain sensitive information. The project has been evaluated by NSD - Privacy services. Case number 102240. It received approval ("assessed with conditions") by NSD on 27/02/2019. Personal data is linked with codes and deleted when the survey is completed. The project complies with institutional guidelines or other ethical regulations:

- Research ethics guidelines for social sciences, humanities, law, and theology (<https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskaps-jus-og-humaniora/> )
- Research Ethics Act (<https://lovdata.no/dokument/NL/law/2017-04-28-23> )
- UiT Guidelines for privacy in research and student projects ([https://uit.no/forskning/art?dim=179056&p\\_document\\_id=604029](https://uit.no/forskning/art?dim=179056&p_document_id=604029))

The legal basis for data collection is the students' consent. We place particular emphasis on not exerting undue pressure on participation, since the context for the data collection is the students' study situation. The students are informed in writing and verbally that participation is voluntary and that they can withdraw whenever they want.

## **Copyright and property rights**

UiT owns the data.

## **Handling of data**

Responsible for data management: Ellen Nierenberg

## **Metadata**

Guidance to the data is created in a ReadMe file. The file describes:

- contact information
- what the data set is about
- additional structure and rules for naming
- where to find which data = overview

## **Appendix 6: Information literacy knowledge test (TILT)**

## Appendix 6: Information literacy knowledge test

All three versions of the Tromsø Information Literacy Test (TILT) are included here:

- [English \(final\)](#)
- [Norwegian \(final\)](#)
- [Pilot \(in Norwegian\)](#)

There are 21 items in the final IL-knowledge tests, 7 for each of the three core facets of IL:

- items 1-7: evaluating sources
- items 8-14: using/citing sources
- items 15-21: seeking information

The four answer alternatives for each item were displayed to participants in random order.

Students were instructed to read the questions and answer alternatives thoroughly, to consider the alternatives, and then choose the answer that they believe is the most correct.

Surveys were made and distributed on Qualtrics. Scores were calculated in Qualtrics after all questions had been answered. Although students were able to see their scores, they could not see which answers were correct since they were due to take the posttest at a later date.

Introductory texts and demographic questions from the surveys are not included in this appendix.

- ✓ Correct answers are marked with a checkmark in this supplement.

## IL-knowledge test: English

1. What characterizes a scholarly article?

- It is written by a researcher from a college, university or other research institution. (1)
- It is published in a printed, English-language journal. (2)
- It is written in plain language that everyone can understand. (3)
- ✓ It is reviewed by independent experts in the field before being published. (4)

2. You find these two books at the library: *Globalization and its Terrors* and *In Defense of Globalization*. Which of these statements do you think is most correct?

- The authors have the same perspective on globalization. (1)
- Both books can be considered objective sources. (2)
- The book titles show the authors' impartiality. (3)
- ✓ To understand globalization, you may want to read (parts of) both books. (4)

3. Which source is *least* suitable for finding information for your article on appropriate nutrition for long distance running?

- academic blogs about running (1)
- scientific journal articles on diet and endurance (2)
- ✓ the website [www.extremefitness.com](http://www.extremefitness.com) (3)
- books/book chapters on diet and sports (4)

4. Which of the following is *not* necessary in order to critically evaluate a source?

- ✓ understand everything in the text (1)
- have some subject-specific expertise (2)
- think critically (3)
- have some knowledge about statistics and methods (4)

5. “The liberal world order will continue and will actually be strengthened by the many anti-liberal outbursts of our time.” Would you say that this quote is:

- Fact (1)
- ✓ Opinion (2)
- Research results (3)
- Theory (4)

6. All of these criteria are useful for assessing the reliability of a journal article, *except for*:

- ✓ the number of tables with statistics (1)
- the author's qualifications (2)
- if the article is peer-reviewed (3)
- if the journal is scholarly (4)

7. You are writing a paper on Arctic ice melting. Who most likely has the *least* reliable information about this on their website?

- The Intergovernmental Panel on Climate Change (1)
- ✓ World Coal Association (2)
- The World Bank (3)
- International Energy Agency (4)

8. Which reference style should you use in your article if this is not specified in the journal?

- You can mix different reference styles, depending on which is most suitable for the source. (1)
- ✓ It doesn't matter which reference style you use as long as you use it consistently. (2)
- You should always use APA. (3)
- If most of your sources are electronic, you should use Harvard style. (4)

9. What is the most important reason to use sources when writing a paper?

- ✓ to support arguments (1)
- to avoid plagiarism (2)
- to show that you've read the sources (3)
- to satisfy the requirements of the assignment (4)

10. In which of the sentences below you do *not* need to cite the source?

- Nine of ten Norwegians buy Christmas presents on the internet. (1)
- ✓ Albert Einstein was born on March 14, 1879 and died on April 18, 1955. (2)
- Men have poorer memories than women. (3)
- According to Vygotsky's sociocultural learning theory, learning occurs through the use of language and participation in social practice. (4)

11. John uses several paragraphs from his own essay written at another university last year. He does not cite the source. Is this considered plagiarism?

- Yes, because his previously written text is not relevant to this assignment. (1)
- ✓ Yes, because it is not permitted to reuse your own texts unless you cite them. (2)
- No, because his previous text was never published. (3)
- No, because it was written at another university. (4)

12. «According to Steenberg (as cited in Spangen, 2005, p. 59), the publishing year is unimportant bibliographical information for works of fiction.» Which source should be included in the reference list?

- Steenberg (1)
- ✓ Spangen (2)
- both Steenberg and Spangen (3)
- neither Steenberg nor Spangen (4)

13. Is it allowed to use an image from a webpage in your article?

- Yes, as long as the article is not published online. (1)
- ✓ Yes, if you have permission from the image's creator and credit them. (2)
- No, because images are copyrighted and can never be used by others. (3)
- No, because images are not covered by Creative Commons-licenses. (4)

14. What kind of information does *not* need to be cited?

- research data (1)
- ✓ common knowledge (2)
- abstracts (3)
- direct quotes (4)

15. You are writing a paper on inclusion in physical education. Which search gives the best results?

- physical education AND inclusion (1)
- (physical education OR gym) AND inclusion (2)
- "inclusion in physical education" (3)
- ✓ inclusion AND "physical education" (4)

16. The most effective way to follow the newest research on genetic testing for autism is to:

- ✓ subscribe to alerts for specific search strings in relevant databases (1)
- search weekly in the library's online catalogue with a specific search string (2)
- follow book releases from publishers specializing in biotechnology (3)
- subscribe to table of contents (ToC) alerts from relevant journals (4)

17. You get only two results for a detailed search in a relevant international scientific database. What should you do to get more results?

- include one more word in the search string (1)
- ✓ exclude one of the search words (2)
- limit the search by year (3)
- search in a language other than English (4)

18. If you wrote comp\* in the search field in a database such as Web of Science, which of the following results would you get?

- accompany (1)
- chomp (2)
- incomprehensive (3)
- ✓ competence (4)

19. In which case is it *not* recommended to use OR between words in a search string?

- when the words are synonyms (1)
- ✓ when you want to get results for just one of the words (2)
- when you want to get results for either of the words (3)
- between a word written in two different languages (4)

20. Which statement is correct?

- When I search correctly in Google, all of the search results are reliable sources. (1)
- ✓ Google Scholar is the academic version of Google, and most search results are scholarly literature. (2)
- I can locate all documents in the library's databases by searching in Google Scholar. (3)
- I get more search results for commercial webpages in Google Scholar than in Google. (4)



21. You are looking for scholarly literature for your paper. If you do not get sufficient results the first time you search, what should you do?

- change the topic of your paper (1)
- ✓ try to find better search terms and search again (2)
- search from another computer (3)
- search only for phrases instead of single words (4)

## IL-knowledge test: Norwegian

1. Hva kjennetegner en vitenskapelig artikkel?

- Den er skrevet av en forsker fra en høyskole, et universitet eller en annen forskningsinstitusjon. (1)
- Den er publisert i et trykt, engelskspråklig tidsskrift. (2)
- Den er skrevet på et allment språk som alle kan forstå. (3)
- ✓ Den blir vurdert av uavhengige eksperter på fagområdet før den blir publisert. (4)

2. Du finner disse to bøkene på biblioteket: *Globalization and its terrors* og *In defense of globalization*. Hvilken av disse påstandene tror du er mest riktig?

- Forfatterne har det samme perspektivet om globalisering (1)
- Begge bøkene kan regnes som objektive kilder (2)
- Boktitlene viser forfatterens upartiskhet (3)
- ✓ For å forstå globalisering kan det være lurt å lese (deler av) begge bøkene (4)

3. Hvilken kilde er *minst* egnet for å finne informasjon til din bacheloroppgave om riktig kosthold for langdistanseløp?

- faglige blogger om løping (1)
- vitenskapelige tidsskriftartikler om kosthold og utholdenhet (2)
- ✓ nettsiden [www.extremefitness.com](http://www.extremefitness.com) (3)
- bøker/bokkapitler om kosthold og idrett (4)

4. Hvilke av punktene nedenfor er *ikke* nødvendig for å kunne kritisk vurdere en kilde?

- ✓ forstå alt i teksten (1)
- ha noe fagspesifikk ekspertise (2)
- tenke kritisk (3)
- ha noe kompetanse innen statistikk og metode (4)

5. “Den liberale verdensorden vil bestå og vil til og med kunne komme styrket ut av vår tids mange antiliberale utbrudd.” Vil du si at dette sitatet er:

- Fakta (1)
- ✓ Meninger (2)
- Forskningsresultater (3)
- Teorier (4)

6. Alle disse kriteriene er nyttige for å vurdere påliteligheten av en tidsskriftartikkel, **bortsett fra**:

- ✓ Antall tabeller med statistikk (1)
- Forfatterens kvalifikasjoner (2)
- Om artikkelen er fagfellevurdert (3)
- Om tidsskriftet er vitenskapelig (4)

7. Du skriver en oppgave om ismelting i Arktis. Hvem har sannsynligvis den **minst** pålitelige informasjonen om dette på sine nettsider?

- The Intergovernmental Panel on Climate Change (1)
- ✓ World Coal Association (2)
- The World Bank (3)
- International Energy Agency (4)

8. Hvilken referansestil skal du bruke hvis den ikke er spesifisert i oppgaveteksten?

- Du kan blande ulike referansestiler, etter hva som passer best til kildetypen. (1)
- ✓ Det spiller liten rolle hvilken stil du bruker, så lenge du bruker den konsekvent. (2)
- Du skal alltid bruke APA. (3)
- Hvis de fleste av dine kilder er elektroniske, skal du bruke Harvardstilen. (4)

9. Hva er den viktigste grunnen til å bruke kilder i en oppgave?

- ✓ For å underbygge argumenter (1)
- For å unngå plagiat (2)
- For å vise at du har lest kildene (3)
- For å tilfredsstille arbeidskrav (4)

10. I hvilken av setningene under trenger du **ikke** oppgi kilden du har brukt?

- Ni av ti nordmenn kjøper julegaver på Internett. (1)
- ✓ Albert Einstein ble født 14. mars 1879 og døde 18. april 1955. (2)
- Menn har dårligere hukommelse enn kvinner. (3)
- Ifølge Vygotskys sosiokulturelle læringsteori skjer læring gjennom bruk av språk og deltakelse i sosial praksis. (4)

11. Regnes dette som plagiat? Jon bruker flere avsnitt fra sin egen oppgave skrevet ved et annet universitet i fjor. Han oppgir ikke kilden.

- Ja, fordi teksten passer ikke til denne oppgaven. (1)
- ✓ Ja, fordi det ikke er lov til å bruke egne tekster om igjen, med mindre man henviser til dem. (2)
- Nei, fordi hans tidligere tekst ikke ble publisert. (3)
- Nei, fordi den var skrevet ved et annet universitet. (4)

12. «Ifølge Steenberg (referert i Spangen, 2005, s. 59), er årstall en uvesentlig opplysning når det gjelder skjønnlitterære bøker.» Hvilken kilde skal oppgis i referanselista?

- Steenberg (1)
- ✓ Spangen (2)
- Både Steenberg og Spangen (3)
- Verken Steenberg eller Spangen (4)

13. Er det lovlig å bruke bilder fra nettsider i en bacheloroppgave?

- Ja, så lenge bacheloroppgaven ikke publiseres på nettet. (1)
- ✓ Ja, hvis du har tillatelse fra opphaver og krediterer vedkommende. (2)
- Nei, fordi bilder er beskyttet av opphavsrett og kan aldri brukes av andre. (3)
- Nei, fordi bilder er ikke dekket av Creative Commons-lisenser. (4)

14. Hvilken type informasjon krever *ikke* kildehenvisning?

- Forskningsdata (1)
- ✓ Allmennkunnskap (2)
- Sammendrag (3)
- Direkte sitat (4)

15. Du skriver en oppgave om tilpasset opplæring i kroppsøving. Hvilket søk gir de beste resultatene?

- tilpasset opplæring AND kroppsøving (1)
- tilpasset opplæring AND (kroppsøving OR physical education) (2)
- "tilpasset opplæring i kroppsøving" (3)
- ✓ "tilpasset opplæring" AND kroppsøving (4)

16. For å følge med den nyeste forskningen om gentesting for autisme, er det mest effektivt å:

- ✓ abonnere på varsler (alerts) for et spesifikt søk i relevante databaser (1)
- søke i bibliotekbasen Oria på det samme spesifikke søk hver uke (2)
- følge med på bokutgivelser fra et forlag som spesialiserer seg i bioteknologi (3)
- abonnere på "TOCs" (table of contents/inneholdsfortegnelser) varsler fra relevante tidsskrifter (4)

17. Du får kun to treff på et detaljert søk i en relevant internasjonal fagdatabase. Hva gjør du for å få flere treff?

- skriv ett ord til i søket (1)
- ✓ fjern ett av søkeordene (2)
- avgrens søket på årstall (3)
- søk på norsk (4)

18. Når du skriver *barn\** i søkefeltet i bibliotekbasen Oria, hvilket ord kan du få treff på?

- enebarn (1)
- låve (2)
- baryton (3)
- ✓ barnehage (4)

19. I hvilket tilfelle er det *ikke* anbefalt å bruke **OR** mellom to ord i et søk?

- når ordene er synonymer (1)
- ✓ når du vil få treff på bare ett av de to ordene (2)
- når du vil få treff på ett av ordene eller begge to (3)
- når det er samme ord fra to forskjellige språk (4)

20. Google Scholar er den akademiske versjonen av Google. Hvilket utsagn om Google Scholar er riktig?

- Alle kildene i trefflista i Google Scholar er fagfelleverderte. (1)
- ✓ Google Scholar gir først og fremst treff på faglige kilder. (2)
- Jeg kan finne alle dokumenter i bibliotekets databaser ved å søke i Google Scholar. (3)
- Jeg får treff på flere offentlige dokumenter i Google Scholar enn i Google (f.eks. Meldinger til Stortinget og NOU-er). (4)

21. Du leter etter faglige kilder til din oppgave. Hvis du ikke får gode nok treff første gang du søker, hva bør du gjøre?

- bytt problemstilling (1)
- ✓ forsøk å finne bedre søkeord og prøv igjen (2)
- søk kun på engelsk (3)
- søk kun på fraser istedenfor enkelte ord (4)

## IL-knowledge test: Pilot (in Norwegian)

All questions in Pilot test begin with P. For the 17 items that were later selected for use in the final IL-knowledge test, final item numbers are indicated in parenthesis, e.g. P8 (1), P20 (16).

Items are grouped according to the six standards in the *Australian and New Zealand information literacy framework*.

Bundy, A. (Ed.) (2004). *Australian and New Zealand information literacy framework: Principles, standards and practice* (2nd ed.). Retrieved from <http://www.libnet.sh.cn/upload/htmleditor/File/130620025617.pdf>

---

### Standard 1: The information literate person recognises the need for information and determines the nature and extent of the information needed.

P5 Du får en oppgave i studiet om et emne som du kan veldig lite om i utgangspunktet. Hvilken kildetype ville du bruke først for å orientere deg litt om temaet?

- ✓ Leksika, inkludert Wikipedia (1)
- Aviser (2)
- Tidsskrifter (3)
- Bøker (4)

P6 Hvem er den tiltenkte målgruppen for denne artikkelen?

**Tittel:** Atomvåpen truer Nord-Koreas supervulkan.

**Abstract:** Det kan bli ISVINTER i Skandinavia og en total KOLLAPS av klodens matvareproduksjon! Halvparten av Asia kan bli BEGRAVET i aske! Konsekvensene er enorme hvis supervulkanen Paektu, i fjellene i Nord-Korea, vekkes til live. Og forskerne er bekymret.

- Geologer (1)
- ✓ Allmennheten (2)
- Forskere/akademikere (3)
- Politikere (4)

P7 Du er interessert i Freuds opprinnelige tanker om psykoanalyse og hvordan han utviklet sin psykoanalytiske teori. Hvilken kilde er det best å bruke?

- Fitzgerald, M. (2017). Why did Sigmund Freud refuse to see Pierre Janet? Origins of psychoanalysis: Janet, Freud or both? *History of Psychiatry*, 28(3), 358-364. (1)
- Håland, E. (2012). Trenger vi psykoanalysen? *Tidsskrift for Norsk Psykologforening*, 49(12), 1205-1209. (2)
- ✓ Freud, S. (1985). *Psykoanalysen: Slik den var - og slik den ble* (W. Brøgger, Overs.). Oslo: Cappelen. (3)
- Appignanesi, R., Zarate, O. & Wyller, T. (1996). *Freud for begynnere*. Oslo: Bracan. (4)

P8 (1) Hva kjennetegner en vitenskapelig artikkel?

- Den er skrevet av en forsker fra en høyskole, et universitet eller en annen forskningsinstitusjon. (1)
- Den er publisert i et trykt, engelskspråklig tidsskrift. (2)
- Den er skrevet på et allment språk som alle kan forstå. (3)
- ✓ Den blir vurdert av uavhengige eksperter på fagområdet før den blir publisert. (4)

P9 I academia er det som regel best å bruke:

- ✓ primærkilder (1)
- sekundærkilder (2)
- tertiærkilder (3)
- kvartærkilder (4)

P10 Du skriver en bacheloroppgave om endringer i kakaoproduksjon i Ghana. Hvilke kilder bør du hovedsakelig bruke?

- ✓ Vitenskapelige tidsskriftartikler (1)
- Leksika, inkludert Wikipedia (2)
- Nyhetsbrev fra Nestlé (3)
- Nettsteder om turisme i Ghana (4)



P11 Hvilke av disse alternativene er det sikreste kjennetegnet på en vitenskapelig tekst?

- Den er skrevet av en professor (1)
- Den inneholder fotnoter (2)
- ✓ Den viser til annen forskning (3)
- Den er gitt ut på et internasjonalt forlag (4)

P12 «Ifølge Steenberg (referert i Spangen, 2005, s. 59), er årstall en uvesentlig opplysning når det gjelder skjønnlitterære bøker.» I denne setningen er det brukt:

- en primærkilde (1)
- ✓ en sekundærkilde (2)
- en tertiærkilde (3)
- en kvartærkilde (4)

P13 Du skriver en 10-siders oppgave om *kolonialisme*, men finner ut at du må avgrense emnet. Hvilken problemstilling kan være aktuell?

- ✓ Hva var effekten av kolonialisme på jordbruk i Elfenbenskysten? (1)
- Hvilke politiske partier har hatt makt i Kenya siden 2000? (2)
- Hvilke land var en del av Det franske koloniriket? (3)
- Hva er definisjonene av et protektorat og en koloni? (4)

P14 (2) Du finner disse to bøkene på biblioteket: *Globalization and its terrors* og *In defense of globalization*. Hvilken av disse påstandene tror du er mest riktig?

- Forfatterne har det samme perspektivet om globalisering (1)
- Begge bøkene kan regnes som objektive kilder (2)
- Boktitlene viser forfatternes upartiskhet (3)
- ✓ For å forstå globalisering kan det være lurt å lese (deler av) begge bøkene (4)

---

**Standard 2: The information literate person finds needed information effectively and efficiently.**

P15 Du finner ikke fulltekst på de artiklene du trenger for din masteroppgave. Hva kan du gjøre?

- Betale for tilgang (1)
- ✓ Spør en bibliotekar om hvordan du kan få gratis tilgang til fulltekst (2)
- Bruk artiklene du får i fulltekst, selv om de er mindre relevant (3)
- Prøv et nytt søk i en annen søkemotor (4)

P16 Hva bør du gjøre for å finne så godt som alle vitenskapelige artikler om de samiske språkene?

- søk i Google (1)
- søk i bibliotekbasen Oria (2)
- søk i databasen MLA (Modern Language Association) (3)
- ✓ søk i flere språkfaglige databaser (4)

P17 (15) Du skriver en oppgave om barneoppdragelse i Japan. Hvilket søk gir deg de mest relevante resultater? (NB: Teksten ble endret etter piloten slik at kun norske ord ble brukt i søkeeksemplet.)

- (barn OR children) AND Japan (1)
- children AND Japan (2)
- ✓ "child rearing" AND Japan (3)
- "raising children in Japan" (4)

P18 (3) Hvilken kilde er *minst* egnet for å finne informasjon til din bacheloroppgave om riktig kosthold for langdistanseløp?

- faglige blogger om løping (1)
- vitenskapelige tidsskriftartikler om kosthold og utholdenhet (2)
- ✓ nettsiden [www.extremefitness.com](http://www.extremefitness.com) (3)
- bøker/bokkapitler om kosthold og idrett (4)

P19 Du vil finne dokumenter skrevet av Vidkun Quisling i 1940. Hvor skal du lete?

- ✓ Riksarkivet (1)
- Forskningsdata (2)
- bibliotekbasen Oria (3)
- databasen Atekst (4)

P20 (16) For å følge med den nyeste forskningen om gentesting for autisme, er det mest effektivt å:

- ✓ abonnere på varsler (alerts) for et spesifikt søk i relevante databaser (1)
- søke i bibliotekbasen Oria på det samme spesifikke søk hver uke (2)
- følge med på bokutgivelser fra et forlag som spesialiserer seg i bioteknologi (3)
- abonnere på "TOCs" (table of contents/inneholdsfortegnelser) varsler fra relevante tidsskrifter (4)

---

**Standard 3: The information literate person critically evaluates information and the information seeking process.**

P21 Du skriver en bacheloroppgave og leter etter relevante kilder. Hvis du synes du får for få treff i bibliotekbasen Oria, kan du få flere relevante treff ved å søke videre i:

- ✓ Google Scholar (1)
- Google (2)
- databasen Atekst (3)
- innholdsfortegnelser av relevante tidsskrifter (4)

P22 Du skal skrive om bekjempelsen av ebola-viruset. Hvilken av disse kildene er mest relevant, tror du?

- Graber-Stiehl, I. (2018). Africa's silent epidemic. *Nature*, 564(7734), 24-26. (1)
- Priomé, A. (2018). Flygende hunder kan stå bak ebolaepidemi. *Illustrert vitenskap*, (6), 72-77. (2)
- ✓ Schnell, M. (2018). Progress in Ebola virus vaccine development. *The Journal of Infectious Diseases*, 215(12), 1775-1776. (3)
- Moe, I. (2018, 20. desember). Fusk, ebola og vold kan ødelegge Kongos valg. *Aftenposten*, s. 26. (4)

P23 (4) Hvilke av punktene nedenfor er *ikke* nødvendig for å kunne kritisk vurdere en kilde?

- ✓ Forstå alt i teksten (1)
- Ha noe fagspesifikk ekspertise (2)
- Tenke kritisk (3)
- Ha noe kompetanse innen statistikk og metode (4)

P24 (5) “Den liberale verdensorden vil bestå og vil til og med kunne komme styrket ut av vår tids mange antiliberale stormkast og kulingbyger.” Vil du si at dette sitatet er:

- Fakta (1)
- ✓ Meninger (2)
- Forskningsresultater (3)
- Teorier (4)

P25 Hva er hensikten med sitatet under? “Ulven i Norge burde få leve i et reservat. Ulv og beitedyr går dårlig overens. Sauer og lam pines til døde. Det blir funnet lam som er så skadet av ulven at de kun går rundt i ring, mens det skummer ut av munnen på dem.”

- Å presentere ulike synspunkter (1)
- Å opplyse (2)
- ✓ Å vekke følelser (3)
- Å overbevise deg til å bli vegetarianer (4)

P26 For å vurdere om en artikkel om røyking og astma er pålitelig, vil du gjerne finne ut om forfatteren er kvalifisert. Da er det mest nyttig å undersøke:

- Antall treff i Google ved søk på forfatterens navn (1)
- Antall artikler forfatteren har publisert de siste 20 årene (2)
- ✓ Om forfatteren tidligere har publisert vitenskapelige arbeider om helse (3)
- Om forfatteren har en faglig blogg om astma (4)

P27 Avsnittet under er hentet fra en artikkel i et nettmagasin. Basert på dette avsnittet, regner du dette som en pålitelig kilde i en bacheloroppgave? “Det er et grensesnitt (en vekselvirkning) mellom jord-pulsen og menneskenes kollektive bevissthet. Dersom menneskenes vibrasjoner holdes nede (den kollektive virkningen av mange lav-frekvente individer) vil dette kunne bidra til at jord-pulsen holdes nede. (Tradisjonelt har dette funnet sted ved at det er blitt spredd lav-frekvens triggere i form av lidelse og frykt blant menneskemassene.)”

- Ja, fordi det virker sannsynlig (1)
- Ja, fordi det er skrevet med fagterminologi (2)
- Nei, fordi det inneholder for mange parenteser (3)
- ✓ Nei, fordi det ikke finnes henvisninger til vitenskapelige kilder (4)

P28 (6) Alle disse kriteriene er nyttige for å vurdere påliteligheten til en tidsskriftartikkel, *bortsett fra*:

- ✓ Antall tabeller med statistikk (1)
- Forfatterens kvalifikasjoner (2)
- Om artikkelen er fagfellevurdert (3)
- Om tidsskriftet er vitenskapelig (4)

P29 Hvordan skal du evaluere en artikkel som sier at medisin mot depresjon *ikke* virker?

- Spør din psykolog om dette er sant. (1)
- ✓ Vurder troverdigheten av forskningen som forfatteren baserer sin påstand på. (2)
- Sammenlign antall artikler som støtter dette synspunktet med antall artikler som sier at medisiner hjelper mot alvorlig depresjon. (3)
- Vurder alle kildene som forfatteren har oppført i sin referanseliste. (4)

P30 Hva er hovedformålet med kildekritikk?

- ✓ Å kvalitetssikre informasjonen du bruker (1)
- Å få best mulig karakter (2)
- Å spare tid ved innlevering av oppgaver (3)
- Å vise at du har lest mange forskjellige artikler/bøker (4)

P31 Hvorfor pleier vi å anbefale at kilder til akademisk bruk bør være tilnærmet objektive?

- Objektive kilder er til å stole på (1)
- Objektive kilder inneholder fakta (2)
- ✓ Objektive kilder har som formål å frembringe ny kunnskap, snarere enn å påvirke (3)
- Objektive kilder har som formål å påvirke, snarere enn å frembringe ny kunnskap (4)

P32 (17) Du får kun to treff på et detaljert søk i en relevant internasjonal fagdatabase. Hva gjør du for å få flere treff?

- skriv ett ord til i søket (1)
- ✓ fjern ett av søkeordene (2)
- avgrens søket på årstall (3)
- søk på norsk (4)

**Standard 4: The information literate person manages information collected or generated.**

P33 Hva slags kilde er dette?

Bergan, M. A. (2017). *Fiskebiologiske undersøkelser i Balsnesvassdraget på Ørland i 2017* (NINA Rapport 1392). Hentet fra <http://hdl.handle.net/11250/2485755>

- Nettside (1)
- Tidsskriftartikkel (2)
- ✓ Rapport (3)
- Rundskriv (4)

P34 (8) Hvilken referansestil skal du bruke hvis den ikke er spesifisert i oppgaveteksten?

- Du kan blande ulike referansestiler, etter hva som passer best til kildetypen. (1)
- ✓ Det spiller liten rolle hvilken stil du bruker, så lenge du bruker den konsekvent. (2)
- Du skal alltid bruke APA. (3)
- Hvis de fleste av dine kilder er elektroniske, skal du bruke Harvardstilen. (4)

P35 Hvilket utsagn om henvisninger i teksten er *feil*?

- En henvisning skal inneholde nok opplysninger til at leseren skal kunne finne referansen i referanselista. (1)
- En henvisning viser at du har brukt en annen kilde, og at stoffet ikke er ditt eget. (2)
- Det er samsvar mellom henvisningene i teksten og referansene i referanselista. (3)
- ✓ En henvisning skal inneholde nok opplysninger til at leserne skal kunne finne frem til kilden. (4)

P36 Du har søkt i Google Scholar og plukket ut mange relevante artikler fra trefflista. Hva er den mest effektive måten å ta vare på dette utvalget før du lukker nettleseren?

- ✓ bokmerke de relevante artiklene i nettleseren (1)
- skriv ut artiklene du vil bruke (2)
- kopier og lim inn avsnitt som du trenger i oppgaven (3)
- send hele trefflista til deg selv via e-post (4)

**Standard 5: The information literate person applies prior and new information to construct new concepts or create new understandings.**

P37 (9) Hva er den viktigste grunnen til å bruke kilder i en oppgave?

- ✓ For å underbygge argumenter (1)
- For å unngå plagiat (2)
- For å vise at du har lest kildene (3)
- For å tilfredsstille arbeidskrav (4)

P38 Du skal holde et foredrag hvor du gjør rede for innvandring til Norge. Hvor er det larest å hente tall fra?

- ✓ De nyeste tallene fra Statistisk sentralbyrå (1)
- Tall fra Antirasistisk Senter (2)
- Tall fra toppåret 2011 fra Statistisk sentralbyrå (3)
- Den nyeste statistikken fra EU (4)

---

**Standard 6: The information literate person uses information with understanding and acknowledges cultural, ethical, economic, legal, and social issues surrounding the use of information.**

P39 Hvorfor får du tilgang hjemmefra til en artikkel du fant i en database, mens din lillesøster på videregående skole ikke får tilgang til den samme artikkelen?

- fordi hun ikke har koblet seg til det trådløse nettverket hjemme (1)
- fordi hun forsøker å få tilgang via mobilen (2)
- ✓ fordi hun ikke er student ved universitetet og dermed ikke har tilgang til universitetsbibliotekets databaser (3)
- fordi hun ikke har riktig passord til databasen (4)

P40 (7) Du skriver en oppgave om ismelting i Arktis. Hvem har sannsynligvis den *minst* pålitelige informasjonen om dette på sine nettsider?

- The Intergovernmental Panel on Climate Change (1)
- ✓ World Coal Association (2)
- The World Bank (3)
- International Energy Agency (4)

P41 (10) I hvilken av setningene under trenger du *ikke* oppgi kilden du har brukt?

- Ni av ti nordmenn kjøper julegaver på Internett. (1)
- ✓ Albert Einstein ble født 14. mars 1879 og døde 18. april 1955. (2)
- Menn har dårligere hukommelse enn kvinner. (3)
- Ifølge Vygotskys sosiokulturelle læringsteori skjer læring gjennom bruk av språk og deltakelse i sosial praksis. (4)

P42 Når trenger du *ikke* henviser til kildene du har brukt?

- Når du bruker en annens teorier. (1)
- ✓ Når du bruker fakta som lett kan slås opp, f.eks. datoen Titanic gikk ned. (2)
- Når du bruker statistikk fra en forskningsartikkel. (3)
- Når du parafraserer en annens tekst. (4)

P43 Regnes dette som plagiat? Oda låner en oppgave fra en tidligere student og skriver den om med egne ord. Hun oppgir ikke kilden.

- ✓ Ja, fordi dette er å presentere andres arbeid som sitt eget. (1)
- Ja, dette kalles «selvplagiat». (2)
- Nei, fordi den tidligere studentens oppgave ikke ble publisert. (3)
- Nei, fordi Oda brukte egne ord. (4)

P44 (11) Regnes dette som plagiat? Jon bruker flere avsnitt fra sin egen oppgave skrevet ved et annet universitet i fjor. Han oppgir ikke kilden.

- Ja, fordi teksten passer ikke til denne oppgaven. (1)
- ✓ Ja, fordi det ikke er lov til å bruke egne tekster om igjen, med mindre man henviser til egne publikasjoner. (2)
- Nei, fordi hans tidligere tekst ikke ble publisert. (3)
- Nei, fordi den var skrevet ved et annet universitet. (4)

P45 Hvilken er den mest korrekte og lærerike måte å skrive en oppgave som krever at du drøfter et emne?

- Lim inn hele avsnitt fra artikler du har lest (med anførselstegn) og sett dem sammen i en logisk rekkefølge. (1)
- ✓ Parafraser deler av tekster du har lest, og argumenter for/imot deres ulike syn. (2)
- Flett sammen kopierte og parafraserte avsnitt fra tekster du har lest. (3)
- Oppsummer tekster du har lest, med egne ord. (4)



P46 Du har brukt et kapittel fra en bok med redaktør i din oppgave. Hvilken opplysning trenger du *ikke* i referansen?

- Kapitelforfatter(e) (1)
- Kapitlets sidespenn (2)
- ✓ Bokens ISBN (3)
- Redaktørnavn (4)

P47 Når du får en oppgave med formen «Redegjør for ... og drøft deretter ...», hvilken måte å skrive på bør du benytte?

- Direkte sitering (1)
- Plagiaring (2)
- Oppsummering (3)
- ✓ Parafrasering (4)

P48 Hvilken er den beste definisjonen av plagiat?

- ✓ Å presentere andres arbeid som sitt eget. (1)
- Å stjele åndsverk. (2)
- Å kopiere tekst med viten og vilje. (3)
- Å gjengi andres ideer eller konklusjoner. (4)

P49 Hva er forskjellen mellom parafrasering og direkte sitering?

- Parafrasering er å gjengi ordrett, mens direkte sitering er å fortelle innholdet med egne ord. (1)
- Parafrasering er å bytte ut "jeg" med navnet på forfatteren, mens direkte sitering er å ikke forandre på noe. (2)
- ✓ Parafrasering er å fortelle innholdet med egne ord, mens direkte sitering er å gjengi ordrett. (3)
- Det betyr det samme: Å gjengi ordrett. (4)

P50 Hva er en lur måte å jobbe på for å få til å oppsummere fra en kilde med egne ord og dermed unngå plagiat?

- ✓ Bli fortrolig nok med kilden til at du kan legge den bort når du skriver. (1)
- Oversett tekstbitene først til et annet språk, og så tilbake til det språket du skriver på. (2)
- Les teksten høyt for deg selv og skriv fort ned det du husker. (3)
- Ha kilden ved siden av deg og gjengi de aktuelle delene av teksten med noen ordforandringer. (4)

P51 (14) Hvilken type informasjon krever *ikke* kildehenvisning?

- Forskningsdata (1)
- ✓ Allmennkunnskap (2)
- Sammendrag (3)
- Direkte sitat (4)

P52 (12) «Ifølge Steenberg (referert i Spangen, 2005, s. 59), er årstall en uvesentlig opplysning når det gjelder skjønnlitterære bøker.» Hvilken kilde skal oppgis i referanselista?

- Steenberg (1)
- ✓ Spangen (2)
- Både Steenberg og Spangen (3)
- Verken Steenberg eller Spangen (4)

P53 Bør vi ha med sidetall i henvisningen?

- ✓ Ja, hvis det hjelper leseren. (1)
- Ja, men bare ved direkte sitat. (2)
- Nei, aldri. (3)
- Nei, fordi de fleste andre gjør det ikke. (4)

P54 (13) Er det lovlig å bruke bilder fra nettsider i en bacheloroppgave?

- Ja, så lenge bacheloroppgaven ikke publiseres på nettet. (1)
- ✓ Ja, hvis du har tillatelse fra opphaver og krediterer vedkommende. (2)
- Nei, fordi bilder er beskyttet av opphavsrett og kan aldri brukes av andre. (3)
- Nei, fordi bilder er ikke dekket av Creative Commons-lisenser. (4)

## **Appendix 7: Assignment-based measures for assessing information literacy skills**

## Appendix 7: Assignment-based measures for assessing information literacy skills

There are two assignment-based information literacy (IL) measures in this project:

- [Source-evaluation measure](#): Annotated bibliography scoring guide for assessing students' abilities to select and critically evaluate sources
- [Source-use measure](#): Rubric for assessing students' abilities to use/cite sources correctly

Both measures are utilized in mandatory assignments for first- and third-year psychology students at UiT The Arctic University of Norway. The two measures were designed by academic librarians in collaboration with other teaching faculty, and results provide authentic assessments of student skills.

See details and examples of the annotated bibliography and rubric measures below.

### Source-evaluation measure

This quantitative method for assessing students' abilities to critically evaluate sources is a variation of a technique devised by Walton and Hepworth (2012).

For their first-year assignment, students chose 3 academic sources for their term papers and composed references to these sources in a bibliography. For their third-year assignment, a bachelor's thesis, students did the same with 10 sources. For each reference, students wrote an annotation describing why it is a good, academic source, and well suited to answering their research question.

Three randomly chosen annotations from students' bibliographies were analyzed and scored with the following component scores (details below):

1. Quality of the source: ratings of each source on a scale of 0 (not academic) to 3 (academic).
2. Variety of criteria: the number of different, unique criteria for source evaluation stated by the student (e.g. relevancy, accuracy, authority, etc.) in each annotation. Each specific criterion is allocated a score of '1'.
3. Frequency of criteria: the number of instances criteria for source evaluation appear in each annotation. Each criterion is allocated a score of '1'. (This will be different from the variety score if a criterion, e.g. relevancy, is listed more than once in the annotation.)

**Quality.** The definition of an academic source provided to students in their first year is composed of three criteria:

- a) An academic source is written by professionals, for professionals.

- b) It builds on other academic sources, cites sources to substantiate claims and arguments, and provides complete references for these sources.
- c) The text has been the subject of a thorough, professional editorial process and has received constructive criticism before being published.

Scoring - quality:

- 0 points – none of the 3 criteria in the definition above (a, b, c)
- 1 point – 1 of the 3 criteria
- 2 points – 2 of the 3 criteria
- 3 points – all 3 criteria

**Variety/frequency.** Students got points if their annotations included criteria for source evaluation in this list:

- authority – e.g. author qualifications, affiliation
- accuracy
- objectivity – bias, self-interest
- currency – up-to-date
- coverage (sample size, e.g. meta-analysis)
- relevance to research question
- publishing channel – e.g. journal/publisher reputation, editors
- quality control – e.g. peer-review
- references provided
- credibility
- level – intended audience
- purpose – why was it published
- authenticity
- reliability
- cited by others – impact
- limitations of the study
- empirical research, IMRAD structure

Below is an example of an annotation to a source in the bibliography of a student whose topic was the effect of chronic pain on mental health:

*This is a good, academic source because it is written by professionals from several universities, it is also peer reviewed and refers to academic sources. I think it may be relevant to include this because substance abuse is one of the consequences that can occur as a result of medication for chronic pain. (translated from Norwegian)*

**Quality:** Rater gave 3 points – the source was an article from a peer-reviewed journal, with references to academic sources, and therefore fulfilled the three quality criteria.

Variety: Rater gave 4 points - one for each of the following unique criteria:

- authority - “*written by professionals*”
- quality control - “*peer reviewed*”
- references provided - “*refers to academic sources*”
- relevance to research question - “*relevant*”

Frequency: Rater gave 5 points - in addition to the 4 criteria above, an extra point was given because the authority criteria was mentioned twice – “*from several universities*”

See Table 1 for examples of the scoring of four students. Two raters scored each annotated bibliography. Averages of raters’ quality, variety and frequency scores were calculated for each student, and these averages were then used in the statistical analyses.

*Table 1: Two raters’ scoring of four students’ annotated bibliographies, showing component scores for quality (Q), variety (V), and frequency (F). Each student evaluated three sources.*

Rater 1												
Student	Source 1			Source 2			Source 3			Average for 6 sources		
	Q	V	F	Q	V	F	Q	V	F	Q	V	F
A	3	2	2	3	1	1	3	2	3			
B	3	5	6	3	5	5	3	5	5			
C	3	6	6	3	4	4	3	3	3			
D	0	3	4	3	2	4	3	2	3			
Rater 2												
A	1	4	7	1	2	3	0	1	2			
B	3	5	6	3	5	5	3	5	5			
C	3	6	6	3	4	4	3	3	3			
D	0	1	2	0	1	2	0	2	2			
Averages of raters’ quality, variety, and frequency scores for each student												
A										1.84	2.00	3.00
B										3.00	5.00	5.33
C										3.00	4.33	4.33
D										1.00	1.83	2.84

### Source-use measure

One of the goals of the introductory psychology course *Thinking, learning, and writing in higher education*, is to prepare students for academic writing. For their final exam, students write a 6-8 page term paper on a psychology topic of their choice, with a clearly defined research question. They are required to use a minimum of five academic sources and to document their sources in the correct manner using APA-style, both in the text and in the reference list. The assignment is graded on a rubric with seventeen evaluation criteria, five of which pertain to the use of sources (see Table 2 for these five criteria). The rubric-criteria can

be adjusted to other reference styles or to relevant learning outcomes in a particular writing assignment when no specific reference style is required.

*Table 2: Criteria and scoring in Source Use rubric for assessing students' use of sources*

Criteria for assessing use of sources	No	Partially	Yes
A) Are academic sources used to support arguments?	0	0.5	1
B) Are sources cited in the text when necessary?	0	0.5	1
C) Are the in-text citations written in correct APA-style?	0	0.5	1
D) Is the reference list written in correct APA-style?	0	0.5	1
E) Are all in-text citations listed in the reference list, and vice versa?	0	-	1

Note. The style name, in this case APA-style, can be exchanged to match the style used for any given assignment.

For this PhD-project, I also assessed the first 6 pages of third-year students' bachelor's theses using the 5 criteria from this rubric shown in Table 1. These five rubric criteria are designed to evaluate the students' abilities to properly use and cite information sources when writing, and students can receive a maximum of five points.

An example of the scoring for four students using the rubric is shown in Table 3.

*Table 3: Rubric scoring for four students*

Criteria/student	1	2	3	4
<b>A)</b>	1	1	0.5	1
<b>B)</b>	1	0.5	0.5	1
<b>C)</b>	1	1	0.5	1
<b>D)</b>	1	0	0.5	1
<b>E)</b>	1	1	0	0
<b>Total</b>	5.0	3.5	2.0	4.0

For more information about the use of rubrics to authentically assess information literacy skills, see Hoffmann and LaBonte (2012).

## References

- Hoffmann, D. A., & LaBonte, K. (2012). Meeting information literacy outcomes: Partnering with faculty to create effective information literacy assessment. *Journal of Information Literacy*, 6(2), 70-85. <https://doi.org/10.11645/6.2.1615>
- Walton, G., & Hepworth, M. (2012). Using assignment data to analyse a blended information literacy intervention: A quantitative approach. *Journal of Librarianship and Information Science*, 45(1), 53-63. <https://doi.org/10.1177/0961000611434999>



**Appendix 8: Tromsø Interest Questionnaire (TRIQ)**

## **Appendix 8: Tromsø Interest Questionnaire (TRIQ)**

Included in this supplement are two sections of TRIQ:

- information about the four phases of interest and questions regarding these
- interest questionnaire with 26 items in 7 subscales

These are provided for both the Self-chosen object of interest (X) and the Specified object of interest (IL).

Introductory texts and demographic questions from the surveys are not included.

### **Self-chosen Object of Interest (X)**

- [Interest phases X](#)
- [Questionnaire X](#)

### **Specified Object of Interest (information literacy - IL)**

- [Interest phases IL](#)
- [Questionnaire IL](#)

## **Self-chosen Object of Interest (X)**

### **Interest phases X**

So many things can interest us, and that interest can take many forms. In this survey, we are interested in learning more about four different kinds of interest that you have or have had. Your first task will involve finding and writing down examples from your life that fit the description of each of these forms.

First read all four descriptions below, and then go back and write in one example for each from your own life that fits all the criteria of the description. Choose an example that you remember and can describe with reasonable ease. You can fill in the examples in whichever order you like. It is absolutely normal if some of the examples are easier to come up with than others.

When you have an example that you think fits, read through the description one more time in order to be sure that your example fits all the points described. If it doesn't hit on all the points, try to come up with another example. If it does, write briefly what the interest is about. Continue this way until you have written in an appropriate example for every form of interest.

When you are done writing in all your examples, answer the questions about each. You will then be sent to more questions asking about ONE of the examples you have provided. You will find out which of the four it will be when you are done and say you are ready to move on.

### **1 - Triggered situational interest**

This interest of mine is one that I haven't had a particular interest in from before. Nonetheless, something about it catches my attention because it matters to me right then and there.

Something about it can feel good in the moment, but that feeling is relatively brief. It's the situations I find myself in that make me notice this interest at all. Otherwise, I might not give it much notice, really. It doesn't mean much to me in my everyday life, even if I sometimes care about it.

Write down an interest you have that fits this category.

---

Q1 How well does the interest description fit for your interest in the interest you wrote down?

- 1--Very poorly (1)
  - 2 (2)
  - 3 (3)
  - 4 (5)
  - 5 (6)
  - 6 -- Very well (7)
- 

Q2 How easy or difficult was it to think of an appropriate example?

- 1 -- Very easy (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 - Very difficult (6)

## **2 - Maintained situational interest**

This interest of mine is one that I knew a little bit about from before. I probably wouldn't have pursued it had I not been in a certain situation (at a particular place, at a certain time and/or with particular people). I notice that engaging with this interest feels good. I can tell that when I do, I often learn new things. As I learn more about it, and/or can do more with it, I come to see how relevant it is for other things I know or do. I also notice that it helps me to have other people or resources around me that support my development of this interest.

Write down an interest you have that fits this category.

---

Q3 How well does the interest description fit for your interest in the interest you wrote down?

- 1--Very poorly (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

Q4 How easy or difficult was it to think of an appropriate example?

- 1 -- Very easy (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 - Very difficult (6)
- 

### **3 - Emerging personal interest**

This interest of mine is one that I really get into. I like pursuing it whenever I can. I am curious about it, and I often find myself wondering about ways to learn more about it or get better at using it in my own way. With practice and use, I can tell I am getting better at or with it, too. That feels really good, and I notice that it means something to me. I am no expert with or at it, so I still need help with it now and then. Fortunately, I get support from other people or other sources when needed.

Write down an interest you have that fits this category.

---

Q5 How well does the interest description fit for your interest in the interest you wrote down?

- 1--Very poorly (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

Q6 How easy or difficult was it to think of an appropriate example?

- 1 -- Very easy (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 - Very difficult (6)
- 

#### **4 - Well-developed personal interest**

This interest of mine is an important part of who I am. I am so interested in it, that I actively pursue developing it -- even when it is hard or challenging. I don't need much help from others. I often figure it out myself with resources I find on my own. Still, I am open for input or feedback from others when I need it. It feels good being competent with this interest. I can tell that people I look up to seem to respect me with this interest of mine. This interest really means a lot to me.

Write down an interest you have that fits this category.

---

Q7 How well does the interest description fit for your interest in the interest you wrote down?

- 1--Very poorly (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

Q8 How easy or difficult was it to think of an appropriate example?

- 1 -- Very easy (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 - Very difficult (6)
- 

Q9 In which month were you born?

- January, February or March (1 – Skip to Q10)
- April, May or June (2 – Skip to Q11)
- July, August or September (3 – Skip to Q12)
- October, November or December (4 - Skip to Q13)

Q10

For the rest of the questionnaire, answer all your questions in relation to the example you gave for interest form **1 - Triggered situational interest**. Just so you remember, write in a brief summary of what you wrote for that example in the space below. Forever hereafter, that interest will be referred to as **X. Remember that!**

---

Q11

For the rest of the questionnaire, answer all your questions in relation to the example you

gave for interest form **2 - Maintained situational interest**. Just so you remember, write in a brief summary of what you wrote for that example in the space below. Forever hereafter, that interest will be referred to as **X. Remember that!**

---

Q12

For the rest of the questionnaire, answer all your questions in relation to the example you gave for interest form **3 - Emerging personal interest**. Just so you remember, write in a brief summary of what you wrote for that example in the space below. Forever hereafter, that interest will be referred to as **X. Remember that!**

---

Q13

For the rest of the questionnaire, answer all your questions in relation to the example you gave for interest form **4 - Well-developed personal interest**. Just so you remember, write in a brief summary of what you wrote for that example in the space below. Forever hereafter, that interest will be referred to as **X. Remember that!**

---

## **Questionnaire X**

### **GENERAL INTEREST – X**

Next, you will be asked to answer questions about your interest in X. For each statement, indicate as best you can which answer fits you best.



1. How interested are you in X?

- Not at all-- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6-- Very (6)
- 

2. How interested are you in being or becoming a person interested in X?

- Not at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6-- Very (6)
- 

3. I think X is very interesting.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### SITUATIONAL INTEREST – X

4. I am dependent upon **others** for maintaining my interest in X.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

5. Without **others**, I would lose interest in X.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

**POSITIVE AFFECT – X**

How little or much do you experience these feelings when you think about your interest in X?

	Not at all -- 1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 -- Very much (6)
6. Pleasure (feeling good) (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Happiness (blissful) (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Interest (curious) (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Engagement (totally immersed) (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**COMPETENCE LEVEL – X**

10. I am satisfied with what I know about X.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

11. Overall, I feel competent in relation to X.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

12. I am satisfied with what I can do with my interest in X.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### COMPETENCE ASPIRATIONS – X

13. I want to **learn** more **about X**

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

14. I want to develop more X-related **skills**.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

15. I want to be able to do more with X.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### MEANINGFULNESS – X

16. Having an interest in X is very useful for me.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

17. I like having an interest in X.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

18. I am very interested in getting good at X.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

19. My interest in X fits well with values that are important to me.

- Not true at all (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

20. My interest in X is very important to me.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

**SELF-REGULATION and ENGAGEMENT – X**

21. I take initiative, on my own, to learn what I can about X.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

22. I make time to develop my X-related knowledge and skills.

- Not true at all-- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

23. I often have questions I wish to find answers to about how to use X well.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

24. I hang in there when developing my interest in X -- even when it is difficult or challenging.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

25. I often take the initiative to use my interest in X more.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

26. I often come up with questions that I want to find answers to related to what I can do with X.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

**COMMENTS:** If you have any other questions or comments, you can write them below.

---

## Interest phases II

### INTEREST AND INFORMATION – First round

We get into doing things that interest us, though what interests us and how much time we spend on our interests may vary widely from person to person and interest to interest. In this study, we want to learn more about what kind of interest you have in **being or becoming an information literate person** and how you experience that. Before you start telling us, though, we have to clarify what we mean first.

By **information**, we mean all information that you meet in your studies and/or everyday life that can have value for how you understand fields of study, and/or life and the world around you. By **information literate person** we mean a person who actively manages to find, evaluate, and use information relevant for their goals. For you, that means being able to:

- Know when you need information
- Find (and get a hold of) information you need when you need it.
- Evaluate information to determine if it is good enough for your purposes
- Use the information in an effective and ethically sound way.

### INTEREST DESCRIPTIONS for Information Literacy

With the above definitions in mind, carefully read the four following interest descriptions. After you have a good understanding of all four, answer the question under each.

---

#### 1- Triggered Situational Interest

I don't have much interest in being or becoming an information literate person. Nonetheless, something about it catches my attention because it matters to me then and there. Being an information literate person can feel good in the moment, but that feeling is relatively brief. I only feel this in certain situations. Otherwise, I might not give much notice to my competence as a user of information. It doesn't really mean much to me in my everyday life, even if I sometimes care about being or becoming an information literate person.

---



Q1 How well does this description fit for how interested you are in being or becoming a more information literate person?

- Very poorly -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

## **2 – Maintained Situational Interest**

My interest in being or becoming a more information literate person is related to something I knew a little about from before. I probably wouldn't have pursued this interest had I not been in a certain situation (at a particular place, at a certain time, and/or with particular people). I notice that being or becoming a more information literate person feels good. I can tell that when I do, I often learn new things. As I learn more about what it means to be information literate and/or how to use information in a good way, I see more how this is relevant for other things I know or do. I also notice also that it helps me to have people or resources around me that support my development as an information literate person.

---

Q2 How well does this description fit for how interested you are in being or becoming a more information literate person?

- Very poorly -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

## **3 - Emerging Personal Interest**

My interest in being or becoming a more information literate person is something I really get into. I like further developing my information literacy knowledge and skills whenever I can. I

am curious about them and I often find myself wondering about ways to learn more about them or become better at using information-relevant knowledge and skills in my own way. With practice and use, I can tell I am getting better at this, too. That feels really good, and I notice that it means something to me. I am no expert, so I still need help with this every now and then. Fortunately, I get support from other people or resources when needed.

---

Q3 How well does this description fit for how interested you are in being or becoming a more information literate person?

- Very poorly -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very well (6)
- 

#### **4 - Well-Developed Personal Interest**

My interest in being or becoming a more information literate person is an important part of who I am. I am so interested in it that I work actively to pursue developing my knowledge and skills in this area – even when it is hard or challenging. I don't need much help from others. I often figure it out myself with resources I find on my own. Still, I am open for input or feedback from others when I need it. It feels good being competent as an information literate person. I can tell that people I look up to in my life respect me for this interest of mine. Being or becoming an information literate person means a lot to me.

---

Q4 How well does this description fit for how interested you are in being or becoming a more information literate person?

- Very poorly -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very well (6)

After you have answered how well each description above fits for you, go on to the next question.

Q5 Which of these four descriptions do you think best describes your interest in being or becoming an information literate person? It doesn't have to fit you exactly, but be the best of the four descriptions provided.

- I am not interested in being an information literate person. (6)
- Interest description 1 (7)
- Interest description 2 (8)
- Interest description 3 (9)
- Interest description 4 (10)

## **Questionnaire IL**

### **GENERAL INTEREST - IL**

For these questions, you will be asked to answer questions about how you experience being an information literate person. For each statement, choose the answer that you believe describes you best.

---

1. How interested are you in **being** an information literate person?

- Not at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6-- Very much (6)

2. How interested are you in **becoming** an information literate person?

- Not at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6-- Very much (6)
- 

3. I think information literacy is very interesting.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### SITUATIONAL INTEREST – IL

4. I am dependent upon **others** for maintaining my interest in being or becoming an information literate person.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

5. Without **others**, I would lose interest in being or becoming an information literate person.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

**POSITIVE AFFECT - IL**

How little or much do you experience these feelings in relation to being or becoming an information literate person?

	Not at all - 1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 - Very much (6)
6. Pleasure (feeling good) (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Happiness (blissful) (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Interest (curious) (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Engagement (totally immersed) (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**COMPETENCE LEVEL – IL**

10. I am satisfied with my information literacy abilities.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

11. Overall I feel I am an information literate person.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

12. I am satisfied with what I accomplish with my level of information literacy.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### COMPETENCE ASPIRATIONS – IL

13. I want to **learn more about** information literacy.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

14. I want to develop more information literacy skills.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

15. I wish I could do more with information literacy.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### MEANINGFULNESS – IL

16. To be an information literate person is very useful for me.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

17. I like being an information literate person.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

18. I am very interested in being or becoming an information literate person.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

19. Being an information literate person fits with values that are important to me.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

20. For me, being or becoming an information literate person is important.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

#### **SELF-REGULATION and ENGAGEMENT – IL**

21. I take initiative, on my own, to learn what I can in order to be more information literate.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-



22. I make time to develop my information literacy knowledge and skills.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

23. I often have questions I want to find answers to about how to use information well.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

24. I hang in there when developing my information literacy – even when it is hard or challenging.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
- 

25. I often take the initiative to become more information literate.

- Not true at all -- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
  - 6 -- Very true (6)
-

26. I often come up with questions that I want to find answers to related to what I **know about** or **can do with** my information literacy.

- Not true at all -- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 -- Very true (6)

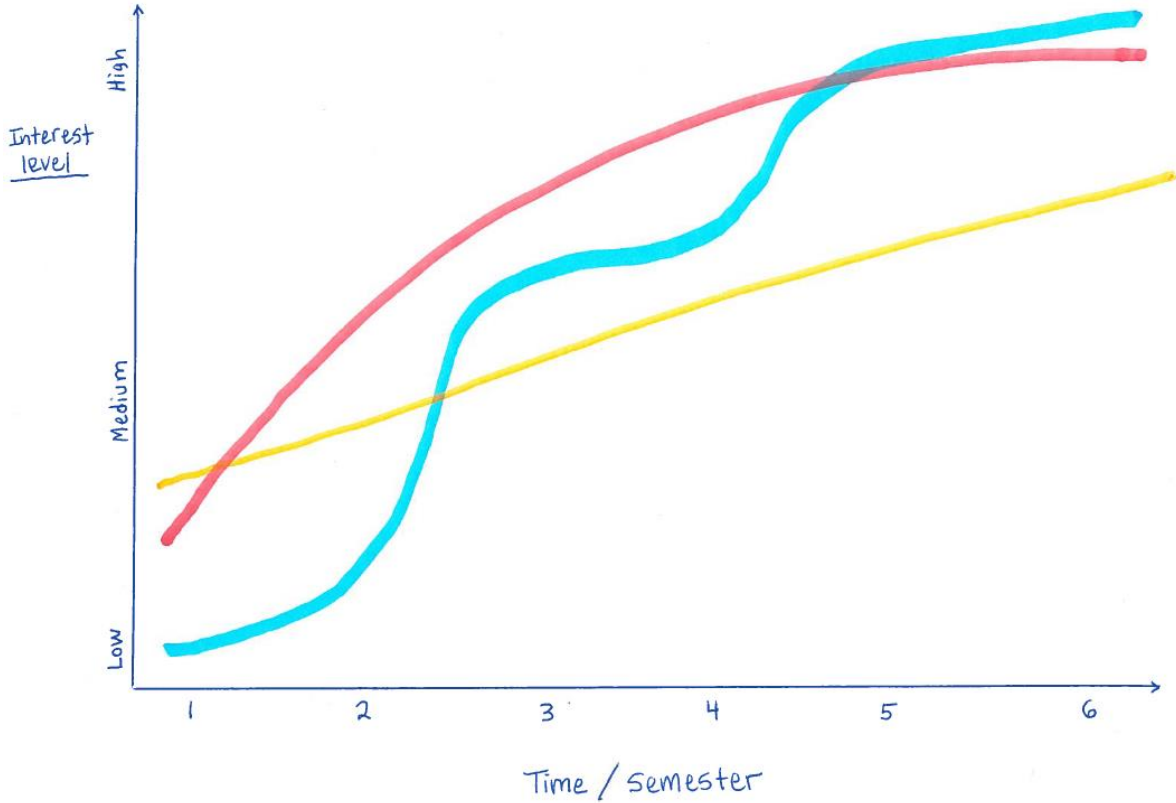
If you have any other questions or comments, you can write them below.

---

**Appendix 9: Interest-o-meter**

# Appendix 9: Interest-o-meter

This interest-o-meter (Paper 4, Figure 2) shows illustrations drawn by the three undergraduates in the second round of individual interviews. It depicts the development of their interest in being or becoming information-literate people over six semesters.



**Appendix 10: Interview guides**

## Appendix 10: Interview guides

### [Focus-group interviews](#)

### [Individual interviews](#)

Focus-group interviews (spring 2021, translated from Norwegian)

Theme	Interview question	Follow-up questions
<b>Attitudes about being an information literate person</b>	<p>What are your thoughts about yourself as an information literate person? How has this developed during your bachelor's studies?</p> <p>How important has it been for you to be critical of info sources; to use sources correctly when writing; to search other places than Google?</p>	Is the way you've planned your literature search for your bachelor's thesis different from the way you would have searched as a new student?
<b>Importance of IL</b>	How important is it for you to be information literate also outside of your studies?	In daily life? At work? As a citizen?
<b>Feelings</b>	Do you experience any special feelings when it comes to your IL, e.g. joy, anxiety, frustration, insecurity, or mastery?	<p>What do you feel when you:</p> <ul style="list-style-type: none"> <li>- search for sources?</li> <li>- evaluate sources?</li> <li>- cite sources?</li> </ul>
<b>Difference between knowing and doing</b>	It is often the case that we don't do what we know is right. You know a lot about IL, but do you actually use what you know in practice when you search for, evaluate, and use sources?	<p>Why not?</p> <p>What's hard about doing what you know is right?</p>
<b>Learning needs</b>	What is your need to learn more about IL in general?	Has this changed since you started your bachelor's degree? Where did you learn about IL? A course? From reading or experience? From friends? Which of these has been the best way to learn IL for you?
<b>Instruction</b>	How satisfied are you with what you have learned about IL at the university?	When have you gotten good help? What do you find most difficult? Is there anything we can do to help you more?
<b>Effort</b>	If you feel the need to learn more, how much effort are you willing to make to become more information literate?	Has this changed since you started your bachelor's degree?

<b>Interest</b>	How interested are you in being or becoming an information literate person?	How has this changed since you started with your bachelor's degree?
<b>Usefulness</b>	How useful is it for you to be information literate?	In which ways?
<b>Values</b>	Would you say that there is a good correlation between being information literate and your other values?	

**Individual interviews** (spring 2022, translated from Norwegian)

<b>Theme</b>	<b>Interview questions</b>	<b>Follow-up questions</b>
<b>Identity</b>	Can you describe what it means to you to be an information literate person?	How well does that describe you?
	How much was being information literate a part of your identity when you started the undergraduate psychology program?	Can you say more about that?
	Has being information literate become part of your identity now? - <i>If yes</i> : Can you describe your transition from being a less to a more information literate person? Were there events, assignments, or anything else that played a role in this transition? - <i>If no</i> : Were you satisfied with who you were regarding your information literacy at the beginning of the program?	Why were these important to your identity as an informational person?  Could you say a little bit about that?
	How did these affect your development as an information literate person? 1. <i>Your knowledge</i> (what you know - cognitive); 2. <i>Your practice</i> (what you do - cognitive); 3. <i>Your feelings</i> (what you feel - emotional); 4. <i>Your environment</i> (social world around you - social)	Has one of these contributed the most to your change of identity (if relevant)?
<b>Transformative learning (TL) (critical reflection)</b>	Some researchers believe that our identity remains the same over time in certain circumstances, while in other circumstances it may evolve as we experience new things that become important to us. Have you experienced anything during your studies that you feel has contributed to changing your identity - who you feel you are as a person?	If so, what did you experience? Tell more about this.
<b>Cognitive, emotional, and social aspects of</b>	Some people believe that learning is more than just what happens cognitively – they believe that learning also involves social and emotional aspects. Do you feel that this is	Can you give me some examples?

<b>learning (identity, TL)</b>	true when it comes to your own learning, i.e., your ways of thinking and acting? Is this the case for you?	
<b>Layers of identity: core, personality layer and preference layers (Illeris)</b>	How much is being information literate part of: <ol style="list-style-type: none"> <li>1. your core identity (self-identity) that is relatively stable throughout life? 'That's who I am.'</li> <li>2. how do you relate to other people, to society, and to important events?</li> <li>3. how you express yourself and feel in everyday situations, and how you react automatically in different situations?</li> </ol>	- Personality: think of a news story in recent years – something that has made a big impression on you. How does your IL affect how you relate to that matter, or how you talk about it with others? - Have these changed with time?
<b>Emotions</b>	Do you experience any special feelings when it comes to your IL, e.g., joy, anxiety, frustration, insecurity, or mastery?	-What do you feel when you are searching for reliable sources, evaluating them, and citing them when you write? - How have these feelings developed over the past three years of study?
<b>Development as an information literate person</b>	How do you assess your own development during the program when it comes to being: <ol style="list-style-type: none"> <li>1. able to search for and find good, relevant sources?</li> <li>2. able to critically evaluate sources of information?</li> <li>3. able to use sources correctly when writing?</li> <li>4. an information literate person, overall?</li> </ol>	Can you say more specifically how the way you (a) find sources, (b) evaluate sources, and (c) use sources has changed since you started higher education?
	If you find that you have become more information literate since you started, was this something happening gradually, or was it a specific episode that triggered it?	If so, what? When did this happen?
<b>Usefulness of IL</b>	Have you needed what you have learned about IL as a student, in daily life, at work, or with health issues?	
	Do you think that being information literate is something that you will need after college, or will it be quickly forgotten when you finish your education?	When might you need it? Does this apply equally to finding, evaluating, and using information sources?
<b>Importance of IL</b>	How has <i>the importance</i> of IL changed for you?	As a student, in daily life, at work, with regard to your health?
<b>Interest</b>	How <i>interested</i> are you in being an information literate person?	How has this interest changed since you started your study program?



	Does your interest in being/becoming an information literate person depend on the situation you are in (e.g., college) or is it <i>independent</i> of the situation (because 'that's just the way I <i>am</i> ', or 'that's how I <i>do</i> things')?	
	If your interest in being or becoming an information literate person is <i>independent</i> of the situation, can you remember approximately when it became that way? When did the interest become more personal, and less dependent on your education?	Was there anything special that happened just then?
<b>Knowing vs. doing</b>	It is often the case that we humans do not do what we know is right. You <i>know</i> a lot about IL, but <i>do you use</i> what you know in practice when searching for, evaluating, and using info sources?	Why, or why not? What's hard about doing what you know is right?
	Now that you have started your bachelor's thesis, how relevant is what you have learned about IL (when you search for, evaluate, and use info)?	
<b>Teaching</b>	How satisfied are you with what you have learned about IL in higher education?	When did you get good help? What do you find the hardest? Is there anything we can do to help you with that?
<b>Learning need</b>	Do you feel a need to learn more about IL in general? If so, what would you like to learn more about? Has this changed since you started the program three years ago?	Where did you learn about IL? A course? Own reading/experience? From friends? Which of these has been the best way to learn IL for you?
<b>Effort</b>	How much effort are you willing to make to become even more information literate? (on a scale, 1 = little effort, 6 = lots of effort).	"If I know myself right, I will make an effort to develop stronger IL skills."
<b>Comments</b>	Is there anything else du would like to say, or personal experiences you would like to share?	

**Appendix 11: Coding of Interviews**

## Appendix 11: Coding of Interviews

Codes, subcodes, and numbers of references in focus group and individual interviews

<b>Codes and subcodes</b>	<b>Ref.</b>	<b>Codes and subcodes (continued)</b>	<b>Ref.</b>
Challenges	12	IL in daily life	
Development		IL after higher education	4
Critical development	15	IL in daily life	11
Development feeling	15	Health IL	1
Development IL attitudes	22	IL in work life	2
Development IL person	16	IL instruction	6
Development IL skills, Do	22	IL term	6
Development IL social	10	Importance IL person	7
Development importance of IL	6	Importance of IL	21
Development knowledge, Know	7	Intent to learn more	3
Development search	30	Interest	
Development source evaluation	27	Interest in being or becoming IL	13
Development source use	21	Situation dependence	20
IL development over time	3	Positive affect	19
Effort learning IL	2	Competency level	8
Feeling		Competency aspirations	10
Feeling about IL	23	Meaningfulness	22
Feeling evaluating sources	15	Lack of interest	3
Feeling finding sources	14	IL interest development	4
Feeling using sources	13	Know vs. Do	8
Feeling social	5	Learning need	1
Feeling, mastery	9	Learning social aspects	12
Feeling combining sources	3	Learning transfer	1
Identity		Metacognition	20
Identity development	12	Motivation	1
Identity IL person	17	Relevance IL	9
Identity source evaluation	5	Risks of not being IL	29
Identity denial	4	Self reflection	6
Identity start	3	Source evaluation	5
Core identity	11	Source searching, finding	1
Personality layer	22	Source use	8
Preference layer	8	Transformation	24
IL attitudes	2	Turn on or off IL	3
		Value of IL	3

