TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

THE ASSESSMENT OF REASONING SKILLS IN NURSING STUDENT SELECTION

Test development and psychometric testing

Jonna Vierula

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1551 | MEDICA – ODONTOLOGICA | TURKU 2021

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

THE ASSESSMENT OF REASONING SKILLS IN NURSING STUDENT SELECTION

Test development and psychometric testing

Jonna Vierula

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1551 | MEDICA – ODONTOLOGICA | TURKU 2021

University of Turku

Faculty of Medicine
Department of Nursing Science
Nursing Science
Doctoral Programme in Nursing Science

Supervised by

Professor Elina Haavisto, RN, PhD
Department of Nursing Science
University of Turku and
Satakunta Central Hospital
Turku, Finland

PhD Maija Hupli, RN
Department of Nursing Science
University of Turku
Turku, Finland

PhD Kirsi Talman, RN
Department of Nursing Science
University of Turku
Turku, Finland

Reviewed by

Professor Tracy Levett-Jones, RN
University of Technology Sydney,
School of Nursing & Midwifery
Sydney, Australia

Docent Hanna-Leena Melender, RNM
Research Unit of Nursing Science and
Health Management
University of Oulu
Oulu, Finland
Director of Nursing Excellence
Hospital District of Helsinki and Uusimaa
Helsinki, Finland

Opponent

Professor Katri Vehviläinen-Julkunen, RN, RM, PhD
Faculty of Health Sciences
Department of Nursing Science
University of Eastern Finland
Kuopio, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-8455-8 (PRINT)
ISBN 978-951-29-8456-5 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku, Finland 2021

“It takes a remarkable person to inspire all of this.”

Richard Gilmore

UNIVERSITY OF TURKU
Faculty of Medicine
Department of Nursing Science
Nursing Science
JONNA VIERULA: The assessment of reasoning skills in nursing student selection – Test development and psychometric testing
Doctoral Dissertation, 155 pp.
Doctoral Programme in Nursing Science
April 2021

ABSTRACT

The purpose of this three-phased study was I) to identify reasoning skills to be assessed in nursing student selection, II) to develop and psychometrically test a Reasoning Skills (ReSki) test for undergraduate nursing student selection, and III) to assess nursing applicants' reasoning skills and factors related to them. The aim was to develop a valid and objective ReSki test for national use in nursing student selection. The ultimate goal was to develop the undergraduate nursing student selection and thus the selection of applicants who are in the future able to make sound decisions thus promoting good nursing care and patient safety.

In phase I, a scoping review (24 original studies) and focus group interviews (16 nursing students and 9 experts) were conducted. Charting, collating, summarising, and inductive content analysis were used to analyse the review data. Deductive and inductive content analysis were used to analyse the focus group data. In phase II, an electronic ReSki test (version 2) was developed (including a pilot-study and expert panels) and psychometrically (including a usability evaluation) tested with a methodological cross-sectional study with 1,056 nursing applicants. The data were analysed using descriptive statistics, correlations, Item Response Theory (IRT) and usability analysis. In phase III, the data (n=1,056) from the psychometric testing (phase II) were used. Descriptive statistics, correlations, and analysis of covariance with Tukey's test in post-hoc multiple group comparisons were used to analyse the data.

The assessment of nursing applicants' reasoning skills involved skills in collecting information, processing information, and identifying the problem and establishing goals. The ReSki test was a valid, usable, and objective assessment method, but the IRT-analysis indicated further improvement to the distractor items is needed for the desired adjustment of the difficulty level. Applicants succeeded best in collecting and processing the information and worst in identifying the problem and establishing goals. Age, gender, previous education, and work experience (only in identifying the problem and establishing goals) were statistically significantly related to the applicants' reasoning skills. The results have implications especially for nursing education and its research.

KEYWORDS: assessment, item response theory, nursing student selection, psychometrics, reasoning skills, test development

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Hoitotieteen laitos
Hoitotiede
JONNA VIERULA: Päätöksentekotaitojen arviointi sairaanhoitajakoulutuksen opiskelijavalinnassa – Testin kehittäminen ja psykometrinen testaus
Väitöskirja, 155 s.
Hoitotieteen tohtoriohjelma
Huhtikuu 2021

TIIVISTELMÄ

Tämän kolmivaiheisen tutkimuksen tarkoituksena oli I) tunnistaa sairaanhoitajakoulutuksen opiskelijavalinnassa arvioitavat päätöksentekotaidot, II) kehittää sairaanhoitajakoulutuksen hakijoiden päätöksentekotaitoja arvioiva Reasoning Skills (ReSki)-testi ja arvioida sen psykometrisia ominaisuuksia, sekä III) arvioida hakijoiden päätöksentekotaitoja ja niihin yhteydessä olevia tekijöitä. Tavoitteena oli kehittää luotettava ja objektiivinen ReSki-testi kansalliseen käyttöön. Lopullinen päämäärä oli kehittää sairaanhoitajakoulutuksen opiskelijavalintaa, mikä edistäisi hyvää hoitoa ja potilasturvallisuutta.

Vaiheessa I toteutettiin kartoittava katsaus (scoping review) (24 tutkimusta) ja fokusryhmähaastattelut (16 sairaanhoitajaopiskelijaa ja 9 asiantuntijaa). Katsausaineisto analysoitiin taulukoimalla, kokoamalla, tiivistämällä ja induktiivisella sisällönanalyysilla, ja fokusryhmäaineisto deduktiivisella ja induktiivisella sisällönanalyysilla. Vaiheessa II kehitettiin sähköinen ReSki-testi (versio 2) (vaihe sisälsi pilottitutkimuksen ja asiantuntijapaneelit). Psykometrinen testaus toteutettiin (sisältäen käytettävyysarvioinnin) poikkileikkaustutkimuksen avulla (1056 sairaanhoitajakoulutuksen hakijaa). Aineisto analysoitiin kuvailevin menetelmin, korrelaatioilla, osiovaste- ja käytettävyysanalyysilla. Vaiheessa III käytettiin samaa aineistoa kuin vaiheessa II (n=1056). Aineisto analysoitiin kuvailevin menetelmin, korrelaatioilla ja kovarianssianalyysilla (Tukeyn testi).

Päätöksentekotaitojen arviointi koostui tiedon keräämisen, tiedon prosessoinnin sekä ongelman tunnistamisen ja tavoitteiden asettamisen taitojen arvioimisesta. ReSki-testi oli luotettava, käytettävä ja objektiivinen arviointimenetelmä, mutta osiovasteanalyysin perusteella harhauttajaväittämiä tulisi muokata testin vaikeustason optimoimiseksi. Parhaiten hakijat menestyivät tiedon keräämisessä ja prosessoinnissa, ja heikoiten ongelman tunnistamisessa ja tavoitteiden asettamisessa. Päätöksentekotaitoja selittivät pääosin ikä, sukupuoli ja koulutustausta sekä osittain työkokemus. Tuloksia voidaan hyödyntää erityisesti hoitotyön koulutuksessa ja sen tutkimuksessa.

AVAINSANAT: arviointi, osiovasteanalyysi, psykometrinen testaus, päätöksentekotaidot, sairaanhoitajakoulutuksen opiskelijavalinta, testin kehittäminen

Table of Contents

Tables and figures

Abbreviations

AERA	American Educational Research Association
APA	American Psychological Association
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CTT	Classical Test Theory
ECTS	European Credit Transfer and Accumulation System
EFN	European Federation of Nurses Associations
ERIC	Education Resources Information Center
EU	European Union
GPA	Grade Point Average
HEI	Higher education institution
HSRT	Health Sciences Reasoning Test
ICC	Item Characteristic Curve
I-CVI	Item-level Content Validity Index
IRT	Item Response Theory
M	Mean
MMI	Multiple Mini Interview
NCLEX-RN	National Council Licensure Examination-Registered Nurse
NCME	National Council on Measurement in Education
NLN	National League for Nursing
NMC	Nursing and Midwifery Council
P-SUS	Positive System Usability Scale
ReSki	Reasoning Skills test
ReSSNE	Reforming Student Selection in Nursing Education (ReSSNE)
RQ	Research question
SAT	Scholastic Aptitude Test
SD	Standard Deviation
SUS	System Usability Scale
SweSAT	Swedish Scholastic Aptitude Test
UAS	University of Applied Sciences
UK	United Kingdom
USA	United States of America
WGCTA	Watson-Glaser Critical Thinking Appraisal
WHO	World Health Organization
2PL	Two parametric logistic model
3PL	Three parametric logistic model

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Vierula, J., Haavisto, E., Hupli, M. & Talman, K. (2020). The assessment of learning skills in nursing student selection: A scoping review. *Assessment & Evaluation in Higher Education*, 45(4): 496–512.

II Vierula, J., Hupli, M., Talman, K. & Haavisto, E. (2020). Identifying reasoning skills for the selection of undergraduate nursing students: a focus group study. *Contemporary Nurse*, 56(2): 120–131.

III Vierula, J., Talman, K., Hupli, M., Laakkonen, E., Engblom, J. & Haavisto, E. (2021). Development and psychometric testing of Reasoning Skills test for nursing student selection: An item response theory approach. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.14799.

IV Vierula, J., Hupli, M., Engblom, J., Laakkonen, E., Talman, K. & Haavisto, E. (2021). Nursing applicants' reasoning skills and factors related to them: A cross-sectional study. *Nurse Education Today*. https://doi.org/10.1016/j.nedt.2021.104890.

1 Introduction

Undergraduate nursing student selection plays a significant role internationally affecting numerous higher education institutions (HEIs) and applicants on an annual basis (Talman et al., 2018a). Nursing is one of the major professions in the world and one of the largest disciplines in HEIs. There are approximately 43.5 million healthcare workers in the world of which nurses and midwives account for nearly 50% (i.e. approximately 20.7 million) (World Health Organization [WHO], 2021a). In Europe, there are an estimated 7.3 million nurses and midwives (WHO, 2021b) of which 72,208 registered nurses actively practice in Finland (Sairaanhoitajat, 2021). In the European Union (EU) member states, there were 121,000 nursing graduates in 2018 (Eurostat, 2020), whereas in Finland, there were 3,372 nursing graduates in 2019 (Vipunen Education Statistics Finland, 2021a). The aforementioned numbers somewhat reflect the magnitude of nursing student selection and the number of applications processed yearly. In Finland, over half of the Finnish higher education applicants are applying to Universities of Applied Sciences (UASs) where nursing is one of the largest degrees and thus there is a high number of applicants each year (Ministry of Education and Culture, 2016; Statistics Finland, 2016; The Rectors' Conference of Finnish Universities of Applied Sciences Arene, 2020). There were 91,890 applications to UAS in the joint application process in Spring 2020 of which 10,641 applied for nursing (programme either Finnish or Swedish) and of which 2,607 were accepted applicants who enrolled to the programme (Vipunen Education Statistics Finland, 2021b).

Selection of the right students to enter nursing programmes is a challenging task for HEIs (Gale et al., 2016) Typically, there are a limited number of study places available and thus the stakes are high for both the applicants and HEIs (Capponi & Mason Barber, 2020; National Council on Measurement in Education [NCME], 2019). According to Talman (2014), student selection can be approached from three perspectives: applicant, HEI and society. From the applicant's perspective, student selection processes should be fair and fluent (Haavisto et al., 2019; Ministry of Education and Culture, 2016; Shulruf et al., 2018; Talman, 2014). From the perspective of HEI, student selection processes should enable the selection of the applicants who will succeed and proceed in their studies fulfilling the competence

requirements of nursing and graduate on time (Schmidt & MacWilliams, 2011; Talman et al., 2018a; Wambuguh et al., 2016). From the societal perspective, student selection processes should enable the selection of applicants who will deliver safe care as future professionals (Shulruf et al., 2018). In addition, student selection processes should aim for a smooth transition from the secondary education to higher education and finally to working life (Ministry of Education and Culture, 2016; Talman, 2014). Furthermore, cost-effective selection practices and the selection of competent and committed future nurses are important for the society, since student withdrawal, nurses' turnover and a lack of competent nurses may cause major societal and economical issues (Andersson et al., 2018; Flinkman, 2014; Hamshire et al., 2019; Ministry of Education and Culture, 2016).

Nursing student selection is an important research topic, since applicants as future students graduate into a registered profession (Ehrenfeld & Tabak, 2000; Haavisto et al., 2019). Furthermore, nursing student selection should be evidence-based (Ministry of Education and Culture, 2016; Haavisto et al., 2019; Hernandez, 2011). However, there is a lack of evidence concerning nursing student selection (Ministry of Education and Culture, 2016; Talman, 2014) and the currently used on-site selection methods in nursing education (Macduff et al., 2016). Moreover, the evidence of effective admission criteria and thus student selection methods is variable as well as the definitions of concepts regarding nursing student selection practices (Hernandez, 2011). It is known that HEIs employ diverse student selection methods resulting in inconsistent selection practices, an absence of evidence-base for most selection methods and a need to justify the best selection practices (Macduff et al., 2016; Talman et al., 2018a; Taylor et al., 2014). In general, nursing education is under scrutiny and accountable for the quality of the education aiming for a positive impact of qualified nurses on patient outcomes (Gale et al., 2016; Royal College of Nursing, 2012). However, nursing student selection is lightly regulated. For example, eligibility for admission in the EU member states, is only regulated by previous education (The European Parliament and the Council of the European Union, 2013). Currently, HEIs are developing their nursing student selection processes aiming to determine exactly what to assess, how to assess it and whom to select (Gale et al., 2016; Haavisto et al., 2019; Macduff et al., 2016; Ministry of Education and Culture, 2017). It has been recommended, that HEIs should use objective, valid, and reliable selection methods (Haavisto et al., 2019; Perkins et al., 2013; Talman et al., 2018a). Furthermore, cost-effectiveness of the selection practices has been discussed recently, because many undergraduate nursing programmes are recruiting more applicants than earlier, but the resources available for education have diminished (Haavisto et al., 2019; Organisation for Economic Co-operation and Development [OECD], 2019; Schmidt & MacWilliams, 2011).

The assessment of cognitive skills is recommended in nursing student selection to select applicants with a wide range of skills to have readiness for both academic and clinical scenarios (McNelis et al., 2010; Talman et al., 2018a; Timer & Clauson, 2011; WHO, 2009). Cognitive skills are vital for nursing students for academic progress (Perkins et al., 2013), and these are emphasised in nursing practice with increasing cognitive requirements (Finnish Institute of Occupational Health (FIOH), 2018). Language, communication, and mathematical skills of nursing applicants have been frequently assessed in nursing student selection (Crouch, 2015; Haavisto et al., 2019; Herrera, 2012; Wolkowitz & Kelley, 2010). However, recent literature recommends the assessment of wider range of cognitive skills in healthcare student selection, namely the assessment of reasoning skills (Haavisto et al., 2019; Lievens et al., 2016; Pitt et al., 2015). Competent professional practice requires not only psychomotor and affective skills but also complex thinking processes because healthcare environments are complex, patients are having multiple health conditions and the use of technology is constantly increasing (Levett-Jones et al., 2010). In addition, the assessed skills of nursing applicants should reflect the requirements of professional education (Schmidt & MacWilliams, 2011; Talman et al., 2018a; Wambuguh et al., 2016). Despite the suggested importance, only few studies have reported the assessment of reasoning skills in nursing student selection and further investigation and operationalisation of the concepts related to reasoning skills in nursing student selection is recommended (Haavisto et al., 2019).

According to Simmons (2010), reasoning skills are cognitive skills focusing on the cognitive (thinking) process preceding decision-making or action (Paper II). In clinical settings, reasoning is defined as a complex cognitive process that uses both formal and informal thinking strategies gathering and analysing patient information, evaluating the significance of this information and weighing alternative actions. Reasoning skills of nurses enable sound decision-making and are crucial for safe nursing practice. (Levett-Jones et al., 2010; Simmons, 2010.) In nursing, decisions are made together with patients, families and multiprofessional teams (Itzhaki et al., 2016; Kae-Hwa & Gyeong-Ju, 2015) in various areas of practice including clinical, administrative and ethical decisions (Johansen & O'Brien, 2016; Kae-Hwa & Gyeong-Ju, 2015; Trobec & Istenic Starcic, 2015). Most importantly, nurses make decisions in direct patient care often under rapid, complex, and uncertain circumstances (Johansen & O'Brien, 2016; Levett-Jones et al., 2010; Simmons, 2010). Nurses make a substantial contribution to healthcare systems and thus decision-making in primary, acute and community care settings (WHO, 2009). Nurses make clinical decisions frequently, and for example, in acute care settings approximately in every ten minutes (Thompson et al., 2013). According to Bucknall (2000), intensive care unit nurses made a decision in every 30 seconds. Nurses' decisions affect care, patient safety, and patient outcomes (Johansen & O'Brien,

2016; Levett-Jones et al., 2010). Nurses' poor reasoning skills in clinical practice, may lead to situations, where nurses are not able to recognise deteriorating patients and they do not intervene appropriately leading to serious consequences (Kim et al., 2015; Levett-Jones et al., 2010). All in all, nurses make decisions constantly, but they are not necessarily aware of the decisions they make (Lauri & Salanterä, 2002). Nurses are required to have capability in making autonomous decisions (Trobec & Istenic Starcic, 2015). The development of nursing students' reasoning skills to cope in clinical situations is an important goal in nursing education (Jessee & Tanner, 2016; Tower et al., 2019; WHO, 2009). Nursing students need reasoning skills from the very beginning of their studies because they face complex clinical decision-making situations in clinical practice with increasing cognitive demands (FIOH, 2018; Levett-Jones et al., 2010).

To sum up, nursing student selection has a vast individual, institutional and societal impact. Selection decisions should be based on objective and valid methods reflecting the requirements of the professional education. However, there is a scarce amount of evidence regarding nursing student selection and the justification of the best selection practices. Furthermore, reasoning skills have been highlighted as crucial skills in nursing and recommended to be assessed already in the selection phase, but the evidence is scarce.

This study focuses on the assessment of reasoning skills in undergraduate nursing student selection. The study was conducted as part of the Reforming Student Selection in Nursing Education (ReSSNE) project that was established nationally in Finland to improve nursing student selection practices (Haavisto et al., 2019). In the ReSSNE project, an evidence-based content and structure for a new joint electronic nursing entrance examination was developed assessing the domains of learning skills (including reasoning skills, language skills and mathematical skills), emotional intelligence and certainty of career choice (Haavisto et al., 2019). After the ReSSNE project, the development of undergraduate nursing student selection (including the assessment of reasoning skills) has continued in the Development Project for Student Selection in Finnish Universities of Applied Sciences. This project was conducted by Finnish UASs and a new national digital entrance exam (UAS Exam) was developed to select students also including nursing education. (Talman et al., 2018b.)

The purpose of this three-phased doctoral dissertation study was I) to identify reasoning skills to be assessed in nursing student selection, II) to develop and psychometrically test a Reasoning Skills (ReSki) test for undergraduate (bachelor level) nursing student selection, and III) to assess nursing applicants' reasoning skills and factors related to them. The aim of the study was to develop a valid and objective ReSki test for national use in nursing student selection. The ultimate goal of the study was to develop the undergraduate nursing student selection and thus the selection of

applicants who are in the future able to make sound decisions thus promoting good nursing care and patient safety.

2 Theoretical Background

In this chapter, the theoretical background of the study is described to provide comprehensive knowledge about the main concepts and central phenomena of this study and to describe the previous knowledge about the study topic. This chapter is divided in three parts. First, nursing student selection in nursing education is described based on previous literature including current legislation and directives relevant to this study. Second, assessment in nursing student selection is described based on previous literature. Third, reasoning skills in nursing education are described based on previous literature and dictionary definitions. Finally, a description of the gaps in the current literature is provided.

2.1 Student selection in nursing education

2.1.1 Nursing education

Nursing education is a link to nursing practice, since the education provides an essential resource to the practice: health professionals (WHO, 2013). However, there is a great variation in the level of degree and how nursing education is organised internationally (WHO, 2009). Nevertheless, all nursing education programmes are recommended to provide education meeting the global standards for the initial education by WHO (2009). Globally, nursing graduates should meet regulatory body standards leading to professional licensure/registration as a nurse (WHO, 2009).

In Europe, undergraduate (bachelor level) nursing education is organised mostly at the higher education level, but variation in the level of degree and implementation of the education exist between countries (Lahtinen et al., 2014). During the last decades, there have been several reforms in the European nursing education attempting to harmonise curricula and degree structures (Kajander-Unkuri et al., 2013). For example, European nursing education has been strongly influenced by the Bologna Process in European Higher Education Area since 1999, which implied further harmonisation and integration of nursing programmes into the higher education system to standardise the nursing education (Collins & Hewer, 2014). Currently, nursing education is based on the EU directives of 2013/55/EU and

2005/36/EC in the member states of the EU (The European Parliament and the Council of the European Union, 2005, 2013). This means that nursing education programmes in EU member states must follow the requirements of the EU directives (Henriksen et al., 2020; The European Parliament and the Council of the European Union, 2005, 2013). According to the requirements of the EU directive (2013/55/EU), training of nurses responsible for general care should comprise a total of at least three years of studies, consisting of at least 4,600 hours of theoretical and clinical training where the theoretical training represents at least one third and clinical training at least one half of the minimum duration of the nursing education. In the EU member states, nursing education is preparing students for a licensed profession (The European Parliament and the Council of the European Union, 2013).

In Finland, undergraduate (bachelor level) nursing education is organised by UASs. Finnish nursing education is part of Finnish higher education consisting of UASs and universities. Specifically, UASs offer a pragmatic education responding to working life needs and emphasising applied research, development, and innovations, whereas universities focus on a scientific research and education. (Ministry of Education and Culture, 2021.) In Finland, nursing education follows the EU directives of 2005/36/EC and 2013/55/EU as other EU member states (The European Parliament and the Council of the European Union, 2005, 2013). Finnish nursing education consists of 210 ECTS (i.e. European Credit Transfer and Accumulation System) taking three and a half years to complete and enabling the application of the right to practice as a licenced professional and using of the occupational title of a healthcare professional (Valvira National Supervisory Authority for Welfare and Health, 2015). In Finland, nurse education covers 210 ECTS, out of which 180 ECTS include core competencies of nurse responsible for general care which are also included in other nursing-based degrees: midwifery, paramedic, public health nursing and diaconal nursing (Silén-Lipponen & Korhonen, 2020). Finnish nursing education is regulated by the Health Care Professionals Act 559/1994, the Universities of Applied Sciences Act 932/2014 and the Government decree on Universities of Applied Sciences 18.12.2014/1129. Furthermore, Finnish nursing education is regulated by the Finnish Ministry of Education and Culture, which agrees together with UASs on the objectives of the education, monitors the education and grants the funding to the UASs. In addition, Finnish nursing education is affected by the healthcare system and its targets and requirements, since UASs are educating professionals based on the needs and requirements of the working life. Although Finnish nursing education is under a scrutiny, Finnish UASs are autonomous meaning that they have a freedom to make independent decisions over their education. (Universities of Applied Sciences Act 932/2014.)

Overall, nursing education should aim for developing competence (Kajander-Unkuri et al., 2020; WHO, 2013). According to WHO (2013), acquisition of nursing

competencies by health professions should target both local and global needs including a culture of critical enquiry and the effective use of information technologies. Furthermore, demands on nurses' competence is increasing due to the changes in healthcare systems, new medical technological solutions, workplace diversity and nurse shortages (WHO, 2016). Developing competence should be continuous process of learning, but it is the duty of formal education (HEIs) to produce the desired quantity and quality of health professionals (Kajander-Unkuri et al., 2020; WHO, 2013). Like in other EU-countries, Finnish nursing education curriculum is competence-based following nationally defined competence areas (Eriksson et al., 2015; Silén-Lipponen & Korhonen, 2020). Although there are national competence requirements and definitions, nursing competences can be considered universal due to the similar requirements of the profession and contents of the higher education studies (European Federation of Nurses Associations [EFN], 2015; WHO, 2009). According to EFN (2015), nursing competence requirements should be based on the EU directive (2013/55/EU) and cover the following main topics: 1) Culture, ethics and values, 2) Health promotion and prevention, guidance and teaching, 3) Decision-making, 4) Communication and teamwork, 5) Research, development and leadership and 6) Nursing care.

2.1.2 Nursing student selection

Undergraduate nursing student selection (i.e. student admission/selection of students/student nurse selection) is a process where nursing programmes select or deselect their students from the pool of applicants (Talman, 2014). Nursing student selection is a path to nursing education whereas the nursing education is a path to working life. Therefore, nursing competence requirements are reflected in nursing student selection, since HEIs seek applicants who will fulfil the competence requirements of nursing during studies (WHO, 2009). In many countries, like Finland and United States of America (USA), there are typically a limited number of new student spaces available, and thus it is important to select students most likely to complete the programme (Capponi & Mason Barber, 2020; National League for Nursing [NLN], 2015; Vipunen Education Statistics Finland, 2021b). In the process of building stronger educational institutions and thus facilitating nursing graduates with competencies relevant for health professional, a key educational policy issue is nursing student selection and the question of how to recruit the right type of students (Frenk et al., 2010; WHO, 2013).

There are several requirements for nursing student selection related to the objectives of selection processes. Nursing student selection should enable the selection of most suitable applicants having both academic and clinical aptitude, who are most likely to succeed in their studies, graduate on time and join the workforce

while delivering appropriate care to healthcare service users (Francis, 2013; Talman et al., 2018a; Wambuguh et al., 2016; WHO, 2009). In addition, the applicants selected should be independent learners, have a realistic perception of nursing and demonstrate the will to serve in healthcare (Haavisto et al., 2019; WHO, 2009). Nursing student selection processes should be transparent, non-discriminatory, cost-effective, and aim for objective assessment using valid and reliable admission tools and thus enabling fair selection practices between applicants (Haavisto et al., 2019; Perkins et al., 2013; Schmidt & MacWilliams, 2011; Talman et al., 2018a; WHO, 2009). All in all, nursing student selection should be evidence-based (Haavisto et al., 2019; Hernandez, 2011). Another challenge for nursing student selection processes is the question of how to select students with diverse backgrounds to increase the gender and ethnic diversity of nursing professionals (Hendricks & Krothe, 2014; Petersen & Lundin, 2007).

Student selection processes vary internationally (Macduff et al., 2016). In general, educational institutions are responsible for organising nursing education and thus responsible for organising nursing student selection (Talman, 2014). Nursing student selection is lightly regulated, and thus the selection processes and admission criteria vary greatly between HEIs and between countries. Therefore, it has been stated that more evidence is needed about the best student selection practices. (Capponi & Mason Barber, 2020; Macduff et al. 2016; Talman et al., 2018a.) Overall, nursing schools use several admission criteria including criteria for the eligibility for admission or minimum entry requirements (such as requirements concerning prior education) and selection methods (such as previous academic achievement and on-site selection) (Capponi & Mason Barber, 2020; Schmidt & MacWilliams, 2011; Talman, 2014; WHO, 2009). According to WHO (2009), nursing schools should have entry requirements that meet national criteria for HEIs including completion of secondary education. In addition, WHO (2009) recommends that nursing schools admit students who meet the institution's health requirements. Furthermore, some HEIs may use open admission policy meaning that a huge number of students are enrolling to a nursing programme meeting only minimum acceptance criteria, which usually leads to a high failure rate during studies and non-enrolments for the second academic year (Shulruf et al., 2011).

In Europe (incl. Finland), nursing student selection has mainly been regulated only in terms of the eligibility for admission and HEIs have been able to independently decide their admission criteria and selection methods (Talman, 2014). Currently, eligibility for admission in the EU-area, is only regulated by previous education. An evidence of completion of general education of 12 years is required from the nursing applicant. (The European Parliament and the Council of the European Union, 2013.) This means, that each EU member state can more specifically state the prior education requirements, admission criteria and the use of

entrance examinations (Haavisto et al., 2019; Lahtinen et al., 2014). For example, countries belonging to EU have been using entry qualifications such as minimum age, good health, or work-experience in the field prior to application (Lahtinen et al., 2014).

In Finland, nursing student selection is organised by UASs and is part of higher education student selection. HEIs are autonomous over their selection processes together with the selection decisions, and thus the selection practices vary between HEIs (Ministry of Education and Culture, 2016). Finnish UASs follow the minimum criteria of previous education set for the EU member states, but otherwise have an autonomous regulation over their student selection practices (Universities of Applied Sciences Act 932/2014). In Finland, the entrance examination results have typically been emphasised as the main admission criterion in nursing student selection like in other higher education disciplines (Ministry of Education and Culture, 2016; Talman, 2014). In Finland, entrance examinations for student selection of UASs are free of charge for the applicants. In addition, Finnish UASs use health requirements for admission stating that an applicant whose state of health or functional capacity makes the applicant incapable of acquitting the practical tasks or clinical practice included in the syllabus cannot bet admitted as a student. (Universities of Applied Sciences Act 932/2014.) Previously, there have been changes in Finnish higher education student selection processes to harmonise the selection practices. For example, Finnish higher education student selection has included the joint application process between HEIs and using a quota for first-time applicants. However, these changes have had only a minor impact on higher education student selection practices. (Ministry of Education and Culture, 2016.) Therefore, higher education student selection (incl. nursing student selection) has been a key development area of Finnish educational policy during the years 2017–2020. According to the latest requirements, selection methods based on the previous academic achievement (i.e. Grade Point Average [GPA] of the secondary education) should be the main admission criteria used, whereas on-site selection methods (i.e. entrance examinations) and other admission criteria (such as open path applications) may be used as complementary criteria, but selection methods used cannot require long preparations from the applicants (Ministry of Education and Culture, 2017).

2.2 Assessment in nursing student selection

2.2.1 Methods used to assess nursing applicants

A variety of student selection methods are used to assess nursing applicants and thus to select nursing students for higher education studies internationally and nationally (Perkins et al., 2013; Schmidt & MacWilliams, 2011; Talman et al., 2018a; Talman

et al., 2018b). Assessment is defined as "the act of judging or deciding the amount, value, quality, or importance of something" (Cambridge Dictionary, 2021). In general, the assessment of students is important for all educational institutions. In student selection, the importance of assessment is emphasised since the methods used for assessment are the basis for selection decisions.

Overall, there are requirements for nursing student selection methods. First, nursing student selection methods need to be able to discriminate applicants and thus to rank them for making the selection decisions. Second, nursing student selection methods should be able to identify the applicants with adequate capacity for the studies. Third, selection methods used should be valid and reliable (Perkins et al., 2013) and they need to be accurately implemented, since applicants to be selected are aiming for a registered profession (Ehrenfeld & Tabak, 2000). In the search of the right students to enter nursing programmes, HEIs are aiming for using selection methods that would be the most effective in predicting future academic performance (Capponi & Mason Barber, 2020; Gale et al., 2016; Hernancez, 2011; Schmidt & MacWilliams, 2011; Wolkowitz & Kelley, 2010). Identifying best predictors of future study success (incl. programme completion) may benefit applicants as future students, nursing faculty, and future healthcare service users by increasing the number of graduates meeting the nursing competence requirements (Capponi & Mason Barber, 2020; Crouch, 2015; Newton et al., 2007).

HEIs use two main methods to select nursing students: previous academic achievement (i.e. preadmission GPA) and on-site selection (i.e. preadmission tests and entrance examinations). All these methods can be considered as part of admission criteria to nursing education.

Previous academic achievement (GPA) is the most used admission criterion and thus the most used selection method to assess nursing applicants internationally (Talman, 2014). Previous academic achievement has included the use of GPA of secondary/college studies and the GPA of preadmit studies (i.e. prerequisite course achievement) (Capponi & Mason Barber, 2020; Talman, 2014). The use of previous academic achievement (GPA) has previously been rare in Finnish higher education (Ministry of Education and Culture, 2016; Talman, 2014). Respectively, previous academic achievement (GPA) is used especially in USA and has been used even as the sole admission criterion to nursing education (Capponi & Mason Barber, 2020). According to the review by Capponi and Mason Barber (2020), research evidence supports the use of previous academic achievement in nursing student selection, since preadmission GPA has been one of the best predictors of nursing student success including programme completion. The prerequisite course achievement (GPA) has included the grades in science related subjects, such as biological sciences, English, and psychology. The prerequisite course achievement (GPA) has predicted early academic achievement, academic success (in separate nursing

courses and in an overall nursing programme) graduation, and National Council Licensure Examination-Registered Nurse (NCLEX-RN) readiness and success in previous studies (Capponi & Mason Barber, 2020; Schmidt & MacWilliams, 2011). However, nursing prerequisite courses differ between the nursing programmes and thus, there is no systematic process at how pre-nursing courses are used in student selection (Hernandez, 2011; Schmidt & MacWilliams, 2011). In addition, some previous studies have also reported that previous academic achievement does not predict future academic performance (Capponi & Mason Barber, 2020; Schmidt & MacWilliams, 2011). According to Stuenkel (2006), previous academic achievement (college GPA) should be combined with other admission criteria. In Stuenkel's (2006) study, previous academic achievement was only a predictor of NCLEX-RN success or failure when combined with other factors (NLN preadmission examination and Scholastic Aptitude Test [SAT] scores). Overall, holistic assessment of nursing applicants is recommended (Macduff et al., 2016; Taylor et al., 2014) and previous academic achievement is suggested to be used together with other admission criteria and not as the sole criteria for selection decisions (Schmidt & MacWilliams, 2011).

On-site selection methods include preadmission tests and entrance examinations, namely, standardised tests and a variety of methods used in entrance examinations, such as literature-based exams, interviews, psychological aptitude tests, admission essays and other types of tests/exams (Capponi & Mason Barber, 2020; Talman et al., 2018a).

Standardised tests refer to tests that are used for monitoring and evaluation purposes and are typically given to large groups of test-takers, and aim to create conditions, questions, scoring procedures and interpretations that are consistent between educational institutions and test-takers (Morris, 2011; Popham, 1999; Wang et al., 2006). Therefore, the benefit of standardised tests is the possibility of objective assessment. Standardised tests are common in USA and often used as preadmission tests for nursing programmes, meaning that the test is performed prior the application, for example in test centres or during the secondary education (Capponi & Mason Barber, 2020). A standardised test is also used in Sweden, where nursing students, together with all higher education applicants, are selected to programmes based on upper secondary school GPA and a standardised test called Swedish Scholastic Aptitude Test (SweSAT) (i.e. Högskoleprovet) (Petersen & Lundin, 2007; Wedman, 2017). Previous literature describes the use of at least 11 standardised tests used in an entry to nursing education: American College Test (ACT), Health Education Systems Inc. (HESI), NLN preadmission test, Nurse Entrance Test (NET), SAT, Test of Essential Academic Skills (TEAS), Watson-Glaser Critical Thinking Appraisal (WGCTA) (Capponi & Mason Barber, 2020), Educational Resources, Inc. (ERI) (Schmidt & MacWilliams, 2011), Health Sciences Reasoning Test (HSRT)

(Pitt et al., 2015), Kaplan admission test (Gartrell et al., 2020) and Nelson-Denny Reading Test (NDRT) (Lajoie, 2013). Standardised tests are effective methods in nursing student selection, since there is an accumulating evidence that standardised tests predict future academic performance (Capponi & Mason Barber, 2020; Gartrell et al., 2020; Haavisto et al., 2019; Hernandez, 2011; Schmidt & MacWilliams, 2011; Stuenkel, 2006; Talman et al., 2018a), although some contradictory results have been reported as well (Schmidt & MacWilliams, 2011). According to Hernandez (2011), standardised tests are more effective than using previous academic achievement (GPA) in selecting nursing students and an integral part of evidence-based nursing student selection processes (Hernandez, 2011). Overall, standardised tests are often used together with other admission criteria, such as previous academic achievement (GPA) (Capponi & Mason Barber, 2020).

Entrance examinations refer to the variety of selection methods used in on-site nursing student selection.

First, a variety of tests/exams have been used in nursing student selection. In Italy, a nationwide entry exam is used to select nursing students among other higher education students (Dante et al., 2011; Lancia et al., 2013). In addition, literature-based exams have been used in nursing student selection (Talman et al., 2018a) and a test measuring emotional intelligence (the Rankein Scale) (Talman et al., 2020).

Second, interviews (individual and/or group interviews) have been one of the most used methods in entrance examinations to assess nursing applicants (Ehrenfeld & Tabak, 2000). Interviews have widely been used in nursing student selection, for example in Finland, in United Kingdom (UK) and in USA (Capponi & Mason Barber, 2020; Djupsjöbacka, 2004; Macduff et al., 2016; Schmidt & MacWilliams, 2011; Talman et al., 2018b). In UK, individual and group interviews are extensively used, since the Nursing and Midwifery Council (NMC) (2011) require that there needs to be a face-to-face contact with students prior to selection to nursing programmes (Macduff et al., 2016). All in all, interviews have been considered as a good complementary selection method for other academic admission criteria (McNelis et al., 2010). The benefit of interviewing has been in the assumed possibility to assess non-academic factors and personal characteristics of nursing applicants (Schmidt & MacWilliams, 2011). However, there has been a lack of research about the effectiveness of interview methods in nursing student selection (Macduff et al., 2016; Schmidt & MacWilliams, 2011; Wolkowitz & Kelley, 2010). Students and faculty have experienced admission interviews positively (Finch et al., 2014), but students and faculty have also had very mixed experiences/opinions on interview processes, for example concerning the validity, reliability and the objectivity of the interview processes (Hendricks & Krothe, 2014; Macduff et al., 2016). Therefore, the challenge in interviewing refers especially to the lack of objective assessment and there has been a lack of substantive evidence about

interviewing (Taylor et al., 2014). Overall, the challenges of interview methods have been recognised in nursing student selection (Macduff et al., 2016). Recently, a more systematic and evidence-based interview method, Multiple Mini Interview (MMI), has been adopted from medical student selection (Eva et al., 2004) and developed to the selection of nursing applicants internationally (Callwood et al., 2020; Gale et al., 2016; Perkins et al., 2013). There is preliminary evidence about the predictive validity of the MMI supporting the potential of MMI to identify the applicants most likely to succeed in clinical practice (Callwood et al., 2020).

Third, psychological aptitude tests have been used in nursing student selection. In Finland, there is a long history (i.e. from the 1940s to the year 2000) of using person specification in nursing student selection, which were dominated by psychological aptitude tests, executed by psychologists (Djupsjöbacka, 2004; Talman, 2014; Talman et al., 2018a). In general, the benefit of psychological aptitude tests has been in the assumed possibility to ensure the suitability of the applicants for the education and thus the selection of safe practitioners (Djupsjöbacka, 2004; Talman, 2014). However, psychological aptitude tests have been criticised during the past twenty years. They have been considered financially burdening and mainly deselecting applicants. Furthermore, the concept of aptitude has not been defined unambiguously. It has been stated that aptitude for the profession is a subjective entity depending on time and place. (Djupsjöbacka, 2004; Ministry of Education and Culture, 2016; Talman, 2014.) Moreover, in a comparison of psychological aptitude tests and literature-based exams, the two selection methods produced very similar results concerning their predictive value in applicants' future nursing studies. Both selection methods predicted nursing students' knowledge, skills, and study success to some extent, but only explained a small proportion of variance. (Talman, 2014; Talman et al., 2018a.) Currently, aptitude is not recommended to be used as a key admission criterion in nursing student selection (Djupsjöbacka, 2004; Ministry of Education and Culture, 2016; Talman, 2014; Talman et al., 2018a).

Fourth, previous studies report the use of admission essays (i.e. description of the personal history and letters of motivation) to assess nursing applicants and to support the selection decisions (Capponi & Mason Barber, 2020; Schmidt & MacWilliams, 2011; Talman et al., 2018a). In addition, personal statements are used in nursing student selection at least in UK (Universities and Colleges Admission Service [UCAS], 2020). However, only one study (Sadler, 2003) has focused on admission essays used in nursing student selection. A statistically significant difference between admission essay scores of "completers" and "non-completers" of the nursing programme was reported. Additionally, "non-completers" tended to write about nursing as external to themselves, based on the thematic analysis. (Sadler, 2003.) Although admission essays may provide helpful information for

selection decision, there is a lack of research on the use and effectiveness of admission essays (Schmidt & MacWilliams, 2011).

Overall, HEIs use a variety of selection methods to assess nursing applicants and combine the selection methods with other admission criteria. However, there is a lack of evidence concerning the relevant skills to be assessed, how to assess them and on what basis to justify the selection practices (Macduff et al., 2016; Perkins et al., 2013; Schmidt & MacWilliams, 2011; Taylor et al., 2014). Many of the selection methods used lack widespread support in previous studies and somewhat conflicting research results may reflect inconsistencies between definition and methodologies used (Schmidt & MacWilliams, 2011). However, evidence suggests that a combination of selection methods and other admission criteria is more effective than any single variable (Schmidt & MacWilliams, 2011). The challenge of identifying the effective, objective and evidence-based nursing student selection methods continues.

2.2.2 Contents used to assess nursing applicants

HEIs are assessing two main types of skills of nursing applicants, namely, cognitive, and non-cognitive skills. Cognitive skills (i.e. learning skills or academic skills) reflect cognitive readiness and academic intelligence of applicants and are needed to succeed in studies, both in theoretical (McNelis et al., 2010) and clinical (Timer & Clauson, 2011) scenarios. Non-cognitive skills reflect the applicant's personal attributes, traits and qualities, temperament, and attitudes (Gale et al., 2016; Talman et al., 2018a). Basically, non-cognitive and cognitive skills are assessed with on-site selection methods. Previous academic achievement (GPA) has been used to assess cognitive skills of the nursing applicants, but in previous literature, there is a lack of detailed description concerning the contents of assessment with GPA (Wolkowitz & Kelley, 2010).

The assessment of cognitive skills (i.e. learning skills or academic skills) in nursing student selection have covered the assessment of decision-making skills (incl. critical thinking and reasoning skills), language and communication skills (incl. reading comprehension, English, writing skills, critical reading, vocabulary and grammar), mathematical skills (incl. mathematics and mathematical reasoning) and science knowledge (incl. biology, anatomy and physiology, chemistry and physics) (Haavisto et al., 2019). Language and communication and mathematical skills are the most frequently assessed skills of nursing applicants (Haavisto et al., 2019; Herrera, 2012; Wolkowitz & Kelley, 2010). Typically, cognitive skills are assessed with previous academic achievement and standardised tests, although MMI has been used to assess both cognitive and non-cognitive skills (Crouch 2015; Haavisto et al., 2019; Herrera 2012; Wolkowitz & Kelley, 2010). In Finland,

cognitive skills have been assessed in nursing student selection, and they are typically described with the concepts of learning skills or academic skills (Djupsjöbacka, 2004; Ministry of Education and Culture, 2016). The assessment of cognitive skills in Finnish health, social, sports, and beauty and cosmetics related UAS programmes' student selection processes (incl. nursing), has comprised the assessment of language skills (Finnish, English, Swedish), reading comprehension, knowledge in basic sciences, ability/readiness in mathematics, logical reasoning ability and problem-solving ability (Talman et al., 2018b). Overall, there is a lack of research on which cognitive skills should be prioritised in the assessment of nursing applicants. Most likely, a comprehensive assessment (i.e. the assessment covers a variety of cognitive skills) of cognitive skills could benefit HEIs. (Haavisto et al., 2019; Schmidt & MacWilliams, 2011.) According to WHO (2009), nursing applicants should have readiness for being independent learners, and thus possess skills in basic science, mathematics, and language of instruction. Overall, cognitive skills have been considered crucial for all higher education applicants, since these skills are needed both in higher education studies and in various fields of working life (Ghanizadeh, 2017; Klegeris et al., 2017). In nursing, the importance of cognitive skills is emphasised. Cognitive requirements for nurses are increasing (FIOH, 2018), cognitive ability is a central nursing competence (Kajander-Unkuri et al., 2013), studying in nursing programme requires extensive reading (Harner, 2014) and nursing students need to fulfil the competence requirements in a relatively short period of time (i.e. 3–4 years of formal education (EFN, 2015). All in all, more research concerning the assessment of nursing applicants' cognitive skills are needed to identify contents of assessment and especially the predictive validity of methods assessing these skills (Talman et al., 2018a).

Non-cognitive skills (i.e. personal attributes, traits and qualities, temperament, and attitudes) are assessed in nursing student selection. HEIs should select applicants who possess characteristics and values (such as compassion, empathy and integrity) needed in nursing, since there are reports of student incivility and a lack of compassion in nursing (Francis, 2013; Gale et al., 2016; Pitt et al., 2014; Waughn et al., 2014). According to WHO (2009), HEIs should select students who demonstrate skills in dealing with clients. The assessment of non-cognitive skills in nursing student selection has covered the assessment of variety of skills, traits or attributes, namely, social skills (incl. interpersonal communication skills such as communication skills, body language, eye contact, expressiveness, interpersonal skills and team-working skills), emotional intelligence (incl. managing and using emotions) and certainty of career choice (incl. nursing awareness and motivation) (Haavisto et al., 2019; Talman et al., 2020). In addition, the assessment of confidence, aptitude for caring, commitment (Macduff et al., 2016), empathy, ethical insights, integrity, initiative, compassion (Callwood et al., 2020; Gale et al., 2016;

Perkins et al., 2013), motivation, kindness, curiosity, honesty, advocacy and respect for privacy, dignity and diversity have been reported (Callwood et al., 2020). Typically, non-cognitive skills are assessed with interviews (Macduff et al., 2016; Perkins et al., 2013; Pitt et al., 2014), but emotional intelligence have been assessed with a test in nursing student selection (Talman et al., 2020). Overall, there is a scarce amount of evidence of the assessment of non-cognitive skills, especially related to the predictive value of nursing student selection methods assessing applicants' non-cognitive skills (Talman et al., 2018a).

Overall, it has been impossible to conclude whether on-site selection methods should prioritise the assessment of cognitive or non-cognitive skills, since the previous research support the assessment of both cognitive and non-cognitive skills (Gale et al., 2016; Talman et al., 2018a). Therefore, the assessment of applicants should focus on a variety of cognitive and non-cognitive skills comprehensively reflecting the requirements of their professional education to support the holistic assessment of nursing applicants (Macduff et al., 2016; Schmidt & MacWilliams, 2011; Talman et al., 2018a; Taylor et al., 2014; Wambuguh et al., 2016). Moreover, more research is needed to identify which cognitive and non-cognitive skills should be assessed in nursing student selection (Macduff et al., 2016; Schmidt & MacWilliams, 2011; Talman et al., 2018a; Taylor et al., 2014).

2.2.3 The development of high-stakes tests

Tests used in student selection are part of high-stakes testing. A test becomes "high-stakes" when the test results lead to serious or important consequences for at least one of the stakeholders. Typically, high-stakes tests are standardised tests and used for admission decisions both for the education and working life or to gain a license to practice (e.g. NCLEX-RN). In student selection, the stakes are high for both the test-taker and HEI: the test-taker is aiming for a study place and the HEI to admit the best applicants. (NCME, 2019; Sackett et al., 2008; Stobart & Eggen, 2012.) In developing a test for student selection purposes, the aim is to construct a test that could discriminate the applicants between the low and high ability test-takers and thus to rank the applicants for making the selection decisions. Therefore, most of the test-takers are not supposed to score the maximum, which challenges the test-designers in constructing questions that not everyone could pass. (Ramsay et al., 2020.)

Typically, high-stakes tests comprise multiple choice questions often having a standard four-alternative structure where only one item is correct (Chiu & Camilli, 2013; Roediger & Marsh, 2005). However, the number of items (i.e. response options) could be more than four, since the test should be constructed in a way that reduces false positives (i.e. a test-taker who did not know the correct answer, could

pass the item by guessing) by decreasing the possibility of guessing (Chiu & Camilli, 2013; DeVellis, 2017). In tests, the incorrect items (i.e. response options not being correct) are called distractor items. Test-designers are aiming for creating functional distractors, referring to incorrect test items that would be tempting to choose by the test-takers, but would be unambiguously incorrect. (Malau-Aduli & Zimitat, 2012.) Furthermore, tests aim to assess test-takers' ability in a specific area, not their ability to guess. However, guessing behaviour is an issue discussed widely in the development of high-stakes tests. The primary challenge of guessing relates to the guessing behaviour of test-takers at lower levels of proficiency. (Chiu & Camilli, 2013; DeVellis, 2017.) If the test is constructed in a way that encourages the test-takers to answer as many questions as possible, regardless of whether they know an answer, the guessing behaviour is likely to increase (Chiu & Camilli, 2013; Rowley & Traub, 1977). However, the possibility of guessing in tests is often called as pseudoguessing, because the guessing behaviour in tests is different from random guessing (Chiu & Camilli, 2013). In some tests, penalty scores are used to avoid test-takers' guessing behaviour. Besides content knowledge, test-taking behaviour, such as risk-taking strategies, motivation, personality factors and test anxiety, are important factors for test-takers' performance and thus guessing behaviour (Burton, 2005; Stenlund et al., 2018). Previous studies have reported that males take more risks when answering test items than females whereas low achievers and women seem to have more test anxiety than males and high achievers (Cassady & Johnson, 2002; Stenlund et al., 2017). In general, it is undesirable for response sets such as risk-taking behaviours or testwiseness to affect students' scores, and thus guessing behaviour should be considered carefully when developing tests (Burton, 2005).

Validity is a crucial in all forms of measurement, but especially in high-stakes tests (Wedman, 2017). Developing a valid test is necessary for the equal treatment of applicants. In high-stakes testing, validity refers to the degree to which all the accumulated evidence supports the intended interpretation of test scores for proposed use, including the evidence base on test content, response processes, internal structure, relations to other variables and consequences of testing. Respectively, the concept of reliability has been used to refer to reliability coefficients of Classical Test Theory (CTT) (such as correlations) and as a more general term to describe the consistency of a test. Reliability is often described independently but has implications for validity. Overall, accumulated evidence should be gathered to develop a valid test. (American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME), 2014.) According to Tavakol et al. (2014), CTT approach has been applied successfully and continuously for many years to assess the psychometric properties of the tests/instruments. However, CTT has some limitations related especially to identification of the item difficulty, and item

discrimination leading to a lack of in-depth interpretation of individual test-items (DeVellis, 2017; Tavakol et al., 2014.) In general, CTT approach emphasises characteristics of a test. For example, reliability in CTT approach is influenced by the correlations of the scale/test items. (DeVellis, 2017.) According to DeVellis (2017), Cronbach's coefficient alpha and therefore reliability, can often be enhanced by redundancy (i.e. increasing the number of items) resulting to a gap in identifying better items rather than the number of items. Furthermore, item parameters (item difficulty and item discrimination) are influenced by different student samples in CTT, which is considered as a limitation. In CTT, the test-taker's ability will appear low if the test item is difficult, while that same test-taker will appear to have a high ability if the questions were easier. Therefore, the Item Response Theory (IRT) was developed based on the characteristics of the items in the test to separate out the characteristics of the test and the sample and to improve the quality of the test items. (DeVellis, 2017; Tavakol et al., 2014; Yang & Kao, 2014.) Currently, the use of IRT approach is increasingly recommended for assessing the validity of measurement scales in nursing, especially when it comes to the validity assessment of tests. Although the importance of IRT has been emphasised in healthcare education, only few psychometricians in nursing education have used IRT methods to create tests that discriminate well at any level of student ability. (Tavakol et al., 2014.)

IRT models are based on mathematical equations explaining the relationship between test-taker ability and the probability of a correct/incorrect item response using a nonlinear monotonic function (Hays et al., 2000; Tavakol et al., 2014). In IRT models, student ability and item parameter values (such as discrimination, difficulty, and pseudoguessing) are transformed mathematically by using natural logarithms into an interval scale resulting in a visual S-shaped logistic curve: the Item Characteristic Curve (ICC) (Paper III, Figure 2). The use of ICCs enables educators and researchers to focus on the assessment of the individual test items, and thus to obtain a greater understanding of the interaction between test-taker ability and item parameters. This helps to monitor and improve the quality of an assessment. (DeVellis, 2017; Dimitrov & Shelestak, 2003; Tavakol et al., 2014.) IRT models may use one or more parameters. The two most well-known IRT models are the one-parameter logistic model (i.e. the Rasch analysis) including the difficulty parameter, and the two-parameter logistic model (2PL) comprising the difficulty and discrimination parameters (Sulis & Toland, 2017). A third parameter, pseudoguessing, is possible to be examined in IRT-models as well (i.e. the three-parameter logistic model [3PL]). The parameters examined are displayed in ICCs. In ICCs, Y-axis is the probability of a correct response whereas the X-axis is the test-taker ability (Paper III, Figure 2). The slope of the ICC describes the discrimination (i.e. the relationship of performance for an item relative to performance on the full test) if an item discriminates amongst weak test-takers. The steeper the slope of the

ICC, the greater the discrimination. The shift of the ICC describes the difficulty of an item referring to the point on the scale where the likelihood of a correct response is 50%. In ICCs, an easy item shifts the curve to the left along the X-axis (test-taker ability), and a difficult item shifts the curve to the right. The pseudoguessing parameter is revealed in the starting point of an ICC along the Y-axis ranging from 0 to 1. Typically, the pseudoguessing is less than 30%. Overall, the higher the starting point of an ICC is, the higher the possibility of guessing is. By examining all the three parameter, useful information is provided about the item parameters and their relationships with the test-taker performance, which may guide test-designers in detecting functional and dysfunctional test-items (DeVellis, 2017; Tavakol et al., 2014.) Overall, IRT approach enables the validity assessment and thus the further development of high-stakes tests to better achieve the aims set for a high-stakes test. In addition, IRT approach views reliability as a fundamental quality of identifying better test items (DeVellis, 2017).

2.3 Reasoning skills in nursing education

2.3.1 The concept of reasoning skills in nursing

Reasoning is a generic and complex concept, typically used synonymously with other concepts describing cognitive (thinking) skills. Dictionaries define reasoning as "a cognitive process directed toward forming conclusions, judgements and inferences from facts or premises" (Webster's Dictionary, 1989), and as "the drawing of inferences or conclusions through the use of reason" (Merriam Webster Dictionary, 2021). Reasoning is a central and important thinking skill that enables making valid inferences, judgements, and reasoned decisions. With reasoning, thinkers can support their conclusions. (Cerbin, 1988; Oxford Cambridge and RSA Examinations [OCR], 2011.) In addition, the concept of reasoning is often used in educational contexts and previous research to describe a wide range of cognitive skills in certain specific areas, such as ethical (Lewis et al., 2019), statistical (Sabbag et al., 2018), spatial, numerical and verbal reasoning (Reinhold, et al., 2020).

Several theories have described the cognitive processes preceding decision-making in various disciplines. In general, these theories are based on a basic assumption that decisions are made either rationally or intuitively. The rational decision-making model describes the cognitive process and its outcome in steps including identifying the problem, generating alternatives, evaluating alternatives, choosing an alternative, implementing the decision, and evaluating the decision's effectiveness (Schoenfeld, 2011). Based on the information processing theory, gathering information, weighing alternative options, and making a final judgement are part of the cognitive process describing how decisions are made (Newell &

Simon, 1972; Simmons, 2010). In nursing literature, the information-processing model (i.e. an analytical model assuming that decision-making follows rational logic) and the intuitive-humanist model (i.e. an intuitive model assuming that decision-making follows intuition and the process is different between novice and expert nurses) have been widely recognised (Banning, 2008; Johansen & O'Brien, 2016). In addition, nursing decision-making theory applied Hammond's cognitive continuum theory (1996) and identified four decision-making models used by nurses: the analytical-systematic model, the analytical-intuitive model, the intuitive-analytical model, and the intuitive-interpretative model. Overall, nursing decision-making was seen (like in Hammond's theory), as a continuum where professionals move flexibly from one model to another depending on the situation, knowledge, and time. (Lauri & Salanterä, 2002.) All the aforementioned theories that are based on the rational assumption of decision-making, suggest that cognitive processes and thus reasoning are required to solve problems, make a decision or to arrive at a solution.

In nursing and other healthcare sciences, reasoning skills refer to cognitive (thinking) skills that are used by healthcare professionals to make decisions in clinical contexts. The concept of clinical reasoning has been used for a long time. Broadly, it refers to cognitive and decision-making processes related to clinical practice. (Higgs & Jones, 1995.) Clinical reasoning in nursing is "a complex cognitive process that uses formal and informal thinking strategies to gather and analyse patient information, evaluate the significance of this information, and weigh alternative actions" (Simmons, 2010). The concept of clinical reasoning is also used in other healthcare disciplines, such as medicine, physical therapy (Huhn et al., 2019) and osteopathy (Grace et al., 2016). However, many healthcare disciplines have struggled in defining clinical reasoning, although it is a vital skill in healthcare practice (Huhn et al., 2019). Compared to nursing, it seems that the concept of clinical reasoning in other healthcare disciplines is more closely related to the concept of diagnostic reasoning (Huhn et al., 2019; Mamede, et al., 2020).

According to Hoffman (2007), clinical reasoning can be defined as "a logical process by which nurses collect cues, process the information, come to an understanding of a patient problem or situation, plan and implement interventions, evaluate outcomes, reflect on and learn from the process" (Levett-Jones et al., 2010). Based on the Hoffman's (2007) original definition, Levett-Jones et al. (2010) developed the Clinical Reasoning Model, which is an educational model representing the clinical reasoning process as a progressive cycle of eight steps: 1) Consider the patient situation, 2) Collect cues/information, 3) Process information, 4) Identify problems/issues, 5) Establish goal/s, 6) Take action, 7) Evaluate outcomes and 8) Reflect on process and new learning. Basically, nurses collect cues, process the information, come to an understanding of a patient's problems, plan, and

implement interventions, evaluate outcomes, and reflect and learn from the process. Clinical reasoning is not a linear process in practice, but possible to be conceptualised as a progressive, step-by-step cycle. (Levett-Jones et al., 2010.) In nursing science, the Clinical Reasoning Model by Levett-Jones et al. (2010) is a well-known and widely used and accepted model that has been used as a theoretical framework in previous studies (Koivisto, 2017; Theobald & Ramsbotham, 2019). In addition, it is noteworthy that reasoning process is always context-dependent and affected by a number of variables, for example cognitive ability, life experience, maturity and skill level within the practice (Simmons, 2010). Reasoning may include inductive, deductive, or abductive approach of thinking (Mirza et al., 2014; Simmons, 2010) and both informal and formal thinking strategies may be used depending on professionals' experience and the situation (Simmons, 2010). Clinical reasoning in nursing comprises knowledge unique to nursing (Simmons, 2010).

In the previous literature, multiple concepts have been used as synonyms with reasoning to describe the cognitive skills that nurses use (Georg et al., 2018; Lauri & Salanterä, 2002; Simmons, 2010). Specifically, the concepts of reasoning, decision-making, problem-solving, clinical judgement, and critical thinking have been used interchangeably to describe how nurses' use their thinking skills in patient care to make decisions (Carbogim et al., 2016; Simmons, 2010). All these concepts have in common that they comprise elements of both the cognitive process and its outcome. However, reasoning emphasises the cognitive processes involved prior to the endpoint whereas the concepts of decision-making, problem-solving and clinical judgement focus on the endpoint of these cognitive processes. (Simmons, 2010.) Moreover, critical thinking is a facilitator of unbiased reasoning and involves both cognitive skills (interpretation, evaluation and inference) and dispositions (attitudes or habits of the mind) (Alfaro-LeFevre, 2013; Ennis, 1985; Facione, 1990; Heijltjes et al., 2015; Hong & Yu, 2017). Critical thinking refers to a broad concept involving knowledge, experiences, dispositions, and intellectual abilities being more than simply learned skills (Carbogim et al., 2016; Facione et al., 1994; Simmons, 2010). Specifically, critical thinking can be defined as "a purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based" (Facione, 1990).

In this study, reasoning is understood as a cognitive skill and as a generic term, referring to a cognitive process leading to forming conclusions, judgements or inferences based on facts or premises (Simmons, 2010; Merriam Webster Dictionary, 2021). Overall, reasoning skills in clinical settings refer to a complex, cognitive process, which uses formal and informal thinking strategies, namely thinking skills, to gather and analyse patient information, evaluate the significance of this information and weigh alternative actions. Clinical reasoning is a logical step-

by-step process involving collecting cues, processing the information, coming to an understanding of a patient problem or situation, planning, and implementing interventions, evaluating outcomes and reflecting on and learn from the process. (Hoffman, 2007; Levett-Jones et al., 2010; Simmons, 2010.) In this study, reasoning skills are understood as generic skills, meaning that such skills do not require nursing specific knowledge in the selection phase. The terminology of clinical reasoning was chosen to this study because clinical reasoning is a well-defined concept/phenomenon in nursing and selected nursing applicants will eventually develop their generic reasoning skills to clinical reasoning skills during their education.

2.3.2 Nursing students' reasoning skills

According to WHO (2009), nursing schools should prepare students with clinical reasoning, problem-solving and critical thinking skills. Specifically, decision-making involving reasoning skills is one of the main competence areas of nursing education nationally (Eriksson et al, 2015; Silén-Lipponen & Korhonen, 2020) and internationally (EFN, 2015; Kajander-Unkuri et al., 2013; Koivisto, 2017; NMC, 2010; WHO, 2009). In practice, a graduating nurse, after three to four years of formal education, should fulfil the competence requirements of nurses to practice safely (WHO, 2009). According to Huang et al. (2018), nursing students develop their reasoning skills in clinical contexts as "an ongoing, interactive, and dynamic process that continuously undergoes adjustments and modifications depending on changes in the clinical context". Decision-making (incl. reasoning skills) is possible to learn with theoretical and practical teaching/learning methods (Koivisto, 2017; Lauri & Salanterä, 2002). However, teaching/learning methods used in nursing education do not always facilitate the development of reasoning skills needed in clinical practice (Koivisto, 2017; Levett-Jones et al., 2010). Based on recent literature, nurse educators often use simulation methods to facilitate the development of nursing students' reasoning skills (Deschênes et al., 2020; Georg et al., 2018; Koivisto, 2017; (Liaw et al., 2018). Furthermore, instruments have been used to assess nursing students' reasoning skills in clinical scenarios, namely the instruments of the Clinical Reasoning Evaluation Simulation Tool (CREST) (Liaw et al., 2018), the Lasater Clinical Judgment Rubric (LCJR) (Georg et al., 2018) and the Script Concordance Test (SCT) (Deschênes et al., 2020). In addition, the Clinical Coaching Interactions Inventory (CCII) was developed to support the development of reasoning skills in clinical practice (Jessee & Tanner, 2016). Based on previous research, a systematic approach is suggested to follow the steps of the reasoning process in clinical scenarios to help nursing students acquire the adequate skills (Cook et al., 2010; (Koivisto et al., 2016; Levett-Jones et al., 2010; Petit dit Dariel et al., 2013; Shipman,

et al., 2012). Specifically, the Clinical Reasoning Model by Levett-Jones et al. (2010) has application for classroom teaching and self-directed learning and can be used to develop computerised learning packages and case methods.

However, previous studies have raised concerns related to nursing students' preparedness to possess reasoning skills and make clinical decisions (Canova et al., 2016; Levett-Jones et al., 2010; Tower et al., 2019; Wolfe, 2017). It has been estimated that over half of all graduating nursing students feel inadequately prepared to make clinical decisions (Harmon & Thompson, 2015). Tower et al. (2019) stated that graduating nursing students' decision-making was inappropriate, whereas Koivisto et al. (2016) stated that in learning clinical reasoning, nursing students learned mainly to take action and collect information compared to the other steps of the reasoning process.

2.3.3 Reasoning skills in nursing student selection

Based on previous literature, reasoning skills in the student selection are generic skills (i.e. they do not involve content specific knowledge), part of learning skills (i.e. cognitive skills or academic skills) and suggested to be taken into consideration in the assessment of nursing applicants (Haavisto et al., 2019; Pitt et al., 2015). Overall, there is scarce amount of studies focusing on the assessment of reasoning skills in an entry phase to nursing education, namely in the student selection. As stated in previous chapter (refer to the chapter 2.2.2), language, communication and mathematical skills are the most frequently assessed skills of nursing applicants meaning that reasoning skills have been assessed less (Haavisto et al., 2019; Herrera, 2012; Wolkowitz & Kelley, 2010). In structuring a new evidence-based entrance examination, Haavisto et al. (2019) recommended the assessment of learning skills, social skills, and certainty of career choice. In detail, the assessment of reasoning skills involving information processing, logical reasoning and problem-solving were suggested together with other learning skills in undergraduate nursing student selection. In addition, there was lack of tests assessing reasoning skills. (Haavisto et al., 2019.) Furthermore, the importance of assessing reasoning skills together with other learning skills has been identified in medicine. For example, the University Clinical Aptitude Test (UCAT, former UKCAT) is used in medical schools to assess clinical aptitude comprising reasoning, decision-making and situational judgement (Lievens et al. 2016; UCAT, 2019). However, more evidence is needed about the contents and methods in nursing student selection. Previously, a challenge of using synonymous concepts in student selection studies have been recognised. (Haavisto et al., 2019; Hernandez, 2011; Macduff et al., 2016; Ministry of Education and Culture, 2016; Taylor et al., 2014.) Therefore, it is recommended that the assessment

of reasoning skills in nursing student selection should be studied and operationalised further (Haavisto et al., 2019).

2.4 Gaps in the current literature

Based on the current literature, undergraduate nursing student selection processes are mainly lightly regulated and thus HEIs use a variety of admission criteria and selection methods to assess nursing applicants. Although the importance of nursing student selection is emphasised in previous literature, there is a lack of evidence about the relevant skills to be assessed, how to assess them and on what basis to justify the nursing student selection practices. Objective assessment is aimed in nursing student selection, but according to the literature, subjective assessment methods have been used. Currently, there is an accumulating evidence of using standardised tests in nursing student selection. The assessment in nursing student selection has comprised both cognitive and non-cognitive factors, but more evidence is needed about the relevant skills to establish the best selection practices. Reasoning skills have been identified being part of learning skills and recommended to be assessed in nursing student selection. However, the literature reports mainly the assessment of other learning skills, such as language and communication and mathematical skills. In addition, there is a lack of tests assessing nursing applicants' reasoning skills. Nursing student selection has been identified being part of high-stakes testing emphasising the importance of validity assessment in test development. High-stakes tests have mainly used CTT approach in psychometric testing, although IRT approach has been recommended as a primary method in high-stakes test development. Overall, the importance of reasoning skills has been emphasised in nursing as a basis of making sound decisions, but the concepts describing the same phenomenon are used interchangeably. Moreover, reasoning skills are core skills of nursing students and registered nurses, but reasoning skills are not emphasised in the selection phase. According to the literature, even graduating nursing students are possibly lacking adequate reasoning skills to make clinical decisions, highlighting the importance of reasoning skills throughout the education, including the selection phase. The current literature supports the need to identify reasoning skills in nursing student selection, operationalise them and study reasoning skills more in-depth.

3 Purpose, aim and research questions

The purpose of this three-phased study (Figure 1) was I) to identify reasoning skills to be assessed in nursing student selection, II) to develop and psychometrically test a Reasoning Skills (ReSki) test for undergraduate (bachelor level) nursing student selection, and III) to assess nursing applicants' reasoning skills and factors related to them. The aim of the study was to develop a valid and objective ReSki test for national use in nursing student selection. The ultimate goal of the study was to develop the undergraduate nursing student selection and thus the selection of applicants who are in the future able to make sound decisions thus promoting good nursing care and patient safety.

The research questions (RQs) addressed were as follows:

I Descriptive phase:

1. What and how reasoning skills are assessed as part of learning skills in undergraduate nursing student selection? (Paper I)
2. What reasoning skills should be assessed in undergraduate nursing student selection? (Paper I, Paper II)

II Test development phase:

3. What are the psychometric properties of the ReSki test for assessing undergraduate nursing applicants' reasoning skills? (Paper III, summary)
4. What is the usability of the ReSki test? (Summary)

III Assessment phase:

5. What is the level of undergraduate nursing applicants' reasoning skills? (Paper IV)
6. What factors are related to nursing applicants' reasoning skills? (Paper IV)

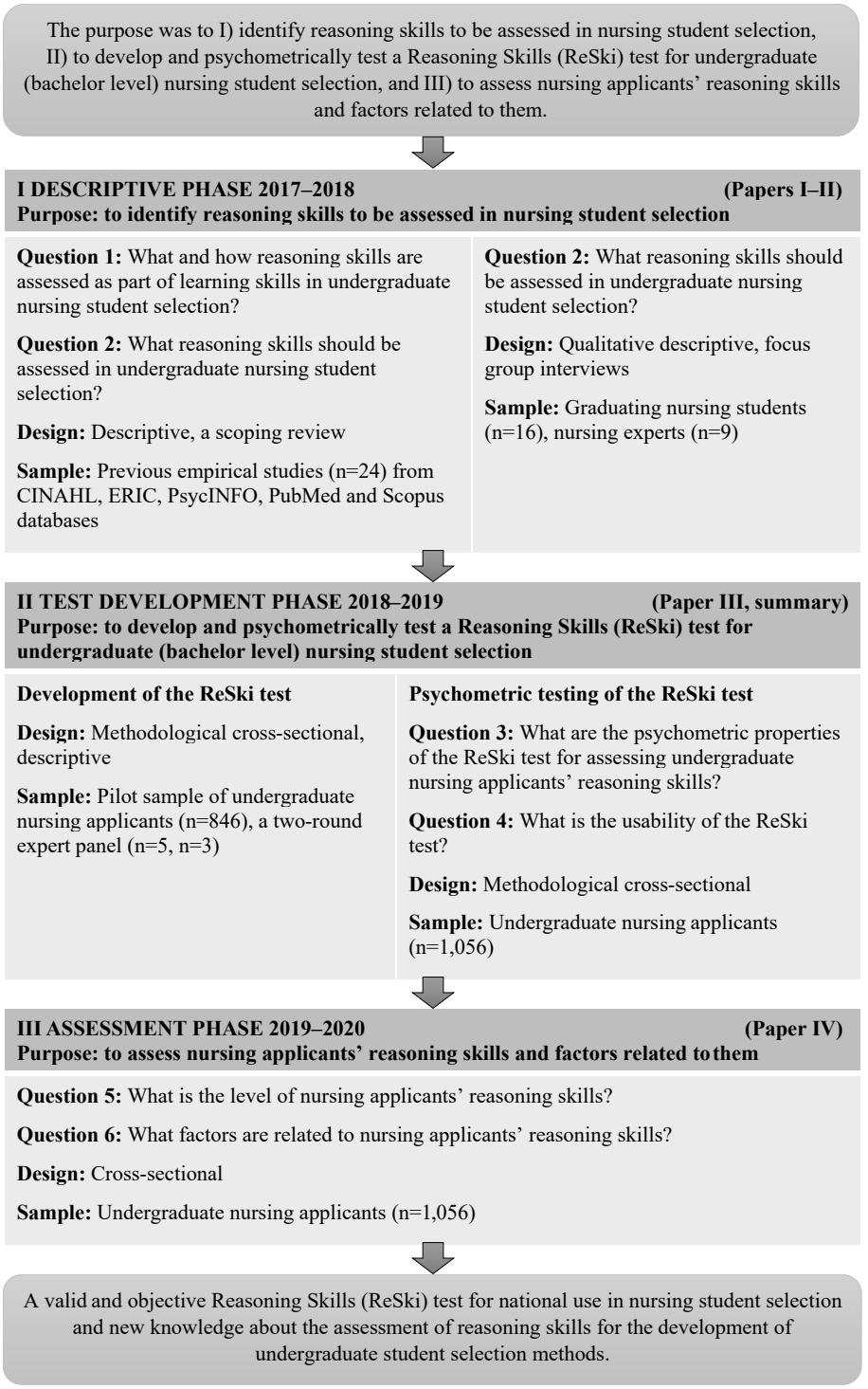

Figure 1. Study phases I–III.

4 Materials and Methods

The aim of the study was to develop a valid and objective ReSki test for national use in nursing student selection. Therefore, the study followed the main principles of test development procedures (AERA, APA & NCME, 2014; DeVellis, 2017). The first version of the ReSki test was developed prior this study. In this study, the second version of the ReSki test was developed and tested.

The ReSki test (version 2) was developed by using mixed methods research (Tashakkori et al., 2020). More specifically, both qualitative and quantitative approaches, and various study designs, samples, settings, data collection and data analysis methods were used to conduct the study (Table 1). The results of the different study phases were integrated to fully understand the research problem and to develop a valid and objective ReSki test for national use in nursing student selection (Moran-Ellis, et al., 2006; Tashakkori et al., 2020). This study was based on a philosophical underpinning of pragmatism (Morgan, 2007; Tashakkori et al., 2020) and the methodological decisions were made based on the test development procedures.

This chapter describes the materials and methods used in this three-phased study to address the RQs 1–6. In addition, the test development process and ethical considerations of the study are described in this chapter.

Table 1. Study phases, designs, and methods.

Phase	Paper	Design	Sample, setting, time	Data collection	Data analysis
I DESCRIPTIVE PHASE 2017–2018	I	Descriptive design: Scoping review	Previous empirical studies (n=24) from CINAHL, ERIC, PsycINFO, PubMed and Scopus databases, 2017–2018	Systematic literature search	Charting, collating, summarising, and inductive content analysis
	II	Qualitative descriptive design	Graduating nursing students (n=16) from two UASs, nursing experts (n=9) from two organisations, 2017–2018	Focus group interviews	Inductive and deductive content analysis
II TEST DEVELOPMENT PHASE 2018–2019	III, summary	Methodological cross-sectional design	Undergraduate nursing applicants (n=846) from six Finnish UASs, 2018	Pilot testing of the ReSki test (version 1)	Descriptive statistics, IRT (2PL)
		Descriptive design	Educators and researchers (1st round: n=5, 2nd round: n=3), 2019	Two-round expert panel	I-CVI
		Methodological cross-sectional design	Undergraduate nursing applicants (n=1,056) from six Finnish UASs, 2019	Testing the psychometric properties of the ReSki test (version 2), P-SUS	Descriptive statistics, correlation coefficients, IRT (3PL), usability analysis
III ASSESSMENT PHASE 2019–2020	IV	Cross-sectional design	Undergraduate nursing applicants (n=1,056) from six Finnish UASs, 2019	ReSki test (version 2), background questionnaire, P-SUS	Descriptive statistics, Analysis of Covariance (Tukey's test)

4.1 Study design, setting and sampling

Phase I

In the first phase of this study, a scoping review (Paper I) and a qualitative descriptive study with focus group interviews (Paper II) were conducted (Table 1).

A scoping review (Paper I) was conducted to map the existing literature and evidence base to identify reasoning skills to be assessed in nursing student selection (Arksey & O'Malley, 2005). The focus of this doctoral dissertation was on the assessment of nursing applicants' reasoning skills, but the preliminary searches indicated the need for a comprehensive analysis of learning skills assessment in nursing student selection, and thus reasoning skills were reviewed as part of other learning skills. Previously, there has been a lack of evidence-based knowledge in relation to nursing student selection practices and the establishment of what exactly to assess (Taylor et al., 2014). A scoping review was chosen as a suitable method to map the contents and methods used in the assessment of nursing applicants' learning skills. The scoping review followed Armstrong et al.'s (2011) five steps of conducting a scoping review: 1) identify the RQs, 2) identify relevant studies, 3) select the studies, 4) chart the data and 5) collate, summarise and report the results. Since scoping reviews do not usually include the assessment of the quality of selected studies, no quality appraisal was conducted (Arksey & O'Malley, 2005; Armstrong et al., 2011). Altogether 24 empirical studies with various study designs were included in the review.

Focus group interviews (Paper II) were conducted to identify reasoning skills to be assessed in nursing student selection for the development of the ReSki test (Doody et al., 2013a). Based on the results of the earlier conducted scoping review (Paper I), reasoning skills were not described on a detailed level enough for the operationalisation of reasoning skills. Focus group interviews including graduating nursing students and experts (nurse educators, managers, and researchers) was considered a suitable method to gather relevant information about a relatively abstract and scarcely studied topic (Doody et al., 2013a). The study informants were purposefully sampled: students from two UASs and experts from two other organisations. Nursing students were recruited via contact persons resulting altogether 16 final year undergraduate nursing students. The decision to include nursing students in the study was based on the informants' recent experience of student selection. Experts were recruited via a contact person or directly by the main researcher resulting altogether nine expert informants. The experts included to the study were assumed to provide relevant insight into the study topic based on their expertise in reasoning skills in nursing. More specifically, the experts were selected based on their research topic and/or teaching expertise in reasoning skills.

Phase II

In the second phase of this study, a ReSki test (version 2) for undergraduate nursing student selection was developed and psychometrically tested (Paper III; Summary) (Table 1). The ReSki test (version 2) was based on the previously developed test version (Paper III, Figure 1). Different study designs were used to develop the ReSki test (version 2) and to test its' psychometric properties (Table 1).

A methodological cross-sectional design was used to pilot test the ReSki test (version 1) to analyse discrimination and the difficulty of the generated items using IRT. The pilot test took place in six Finnish UASs nationwide with 846 undergraduate nursing applicants who took the joint electronic entrance examination of the RESSNE project and consented to the pilot study (Paper III). The participants were purposefully sampled, because they participated in the entrance examinations organised by the UASs who were partners of the ReSSNE project. In the pilot sample, the mean (M) age of the applicants were 26.5 years (Standard Deviation [SD] 8.1, range 18–57), mostly being female (n=738, 87.6%) and 44.1% were first-time applicants.

After pilot testing, a descriptive design with a two-round expert panel was used for content validity evaluation of the ReSki test (Paper III). Six experts who were educators and/or researchers were purposefully sampled and invited to the study. The experts were from one organisation and they had participated to the focus group interviews in the first phase of this study. Five experts participated to the first round. In the second round, three (out of four) of the experts' responses followed the instructions provided and thus were included in the study.

After the pilot testing and content validity evaluations, a methodological cross-sectional design was used for psychometric testing of the developed ReSki test (version 2) (Paper III). Six Finnish UASs who participated to the pilot testing were included in the study. The study included a purposeful sample of 1,056 (N=1,906) undergraduate nursing applicants who took the joint electronic entrance examination and consented to the study.

Phase III

In the third phase of this study, a cross-sectional design was used to assess nursing applicants' reasoning skills and factors related to it (Paper IV) (Table 1). The sample (n=1,056) from the psychometric testing of the ReSki test (version 2) (refer to the description of the psychometric testing in study phase II, page 43) was also used in the third study phase.

4.2 Data collection

Phase I

In the scoping review (Paper I), the data were collected with systematic literature search from five electronic databases (Cumulative Index to Nursing and Allied Health Literature [CINAHL], Education Resources Information Center [ERIC], PubMed, PsycINFO and Scopus). The data search combined relevant Medical Subject Headings and CINAHL headings and several dictionary terms describing learning skills and nursing student selection (more in-depth description in Paper I). The search terms included the concept of reasoning and its synonyms to ensure scoping of studies that comprehensively addressed the assessment of nursing applicants' learning skills including reasoning skills. The search was limited to English or Finnish language, abstract available and publication year 2006–2018 to capture the most recent literature. The identification of the relevant studies was conducted by two researchers in 10 July 2018. The selection of relevant studies was based on the inclusion/exclusion criteria by screening the title, abstract and full text (Aveyard, 2007). Empirical studies, literature reviews and doctoral dissertations reporting assessment of undergraduate nursing applicants' learning skills were included in the review. After the systematic search and identification of relevant studies, the reference lists of the retrieved articles were manually searched for the possible inclusion of relevant articles outside the databases used resulting to the inclusion of two studies. As a result of the literature search, 24 original studies (published during the years 2006–2016) were selected, namely 19 research articles and 5 doctoral dissertations.

In the qualitative descriptive study (Paper II), the data were collected with focus group interviews from two expert groups and two student groups (n=25) (Doody et al., 2013b). The interviews were conducted face-to-face in meeting rooms of the informants' organisations (n=2) in December 2017 and January 2018. The interviews followed a previously piloted interview guide (Vaughn et al. 1996) that included: 1) Introduction, 2) signing the informed consent, 3) filling the short background information questionnaire, 4) a warm-up and clarification of the terms, 5) main question ("What reasoning skills should be assessed in nursing student selection"?) and 6) a wrap-up and closing statements (Paper II). One moderator (the main researcher) conducted the interviews and applied reflective listening and follow-up questions to encourage in-depth and interactive discussion of a study topic (Doody et al., 2013b). The interviews ranged from 60 to 75 minutes. All the interviews were audio-recorded and transcribed verbatim by the main researcher.

Phase II

In the pilot study (Paper III), the data were collected (31 October 2018) with the ReSki test (version 1). The nursing applicants (n=846) performed the test under supervision in computer classrooms as part of the joint electronic entrance examination consisting of the domains of learning skills (including reasoning skills, language skills and mathematical skills), emotional intelligence and certainty of career choice (Haavisto et al., 2019).

After the pilot study, the ReSki test was further developed (Paper III). The data were collected (March 2019) from an expert panel (two rounds) with electronic questionnaires via e-mail for the evaluation of ReSki test question and item (i.e. response option) relevance and clarity (DeVon, et al., 2007; Paper III). In the first round, the questionnaire included the structure and content of the ReSki test, and the experts were invited to evaluate the questions/items on a scale relevant/not relevant and clear/not clear (DeVon et al., 2007; Polit & Beck, 2006). Additionally, the experts were asked to provide comments and suggestions for revisions if necessary. In the second round, the experts evaluated the clarity (clear/not clear) of the revised questions/items and the relevance of the distractor items (i.e. incorrect items of the test) in comparison to the correct items to avoid ambiguity in the correct answers. Experts were able to provide comments and suggestions for revisions.

After the pilot study and content validity evaluation, in the psychometrical testing (Paper III), the data were collected (28 May 2019) with the ReSki test (version 2). Similar to pilot study, the nursing applicants (n=1,056) performed the ReSki test under supervision in computer classrooms as part of the joint electronic entrance examination consisting of the domains of learning skills (including reasoning skills, language skills and mathematical skills), emotional intelligence and certainty of career choice (Haavisto et al., 2019). Furthermore, a Finnish version of The Positive System Usability Scale (P-SUS) (Jokela, 2013) was used as part of the data collection to assess how the applicants as subjective users perceived the usability of the electronic ReSki test for student selection purposes. The usability (i.e. effectiveness, efficiency, and satisfaction) of any tool or system must be viewed for its appropriateness to the certain context (Brooke, 1996; Brooke, 2013).

Phase III

In the cross-sectional study (Paper IV), the data were collected with validated ReSki test (version 2) and P-SUS. The data (n=1,056) from the psychometric testing of the ReSki test (version 2) (refer to the description of the psychometric testing in study phase II, page 45) was also used in the third study phase. In addition, demographic details (age, gender, previous education [high school/vocational school], previous Finnish higher education degree, previous application to nursing studies, study

programme choice and work experience in the field prior to application) were collected with a background questionnaire in the joint electronic entrance examination.

4.3 Instruments

Two instruments, namely ReSki test and P-SUS, were used in this study in phases II and III (Table 1). The ReSki test (version 2) was developed in this study (Figure 2).

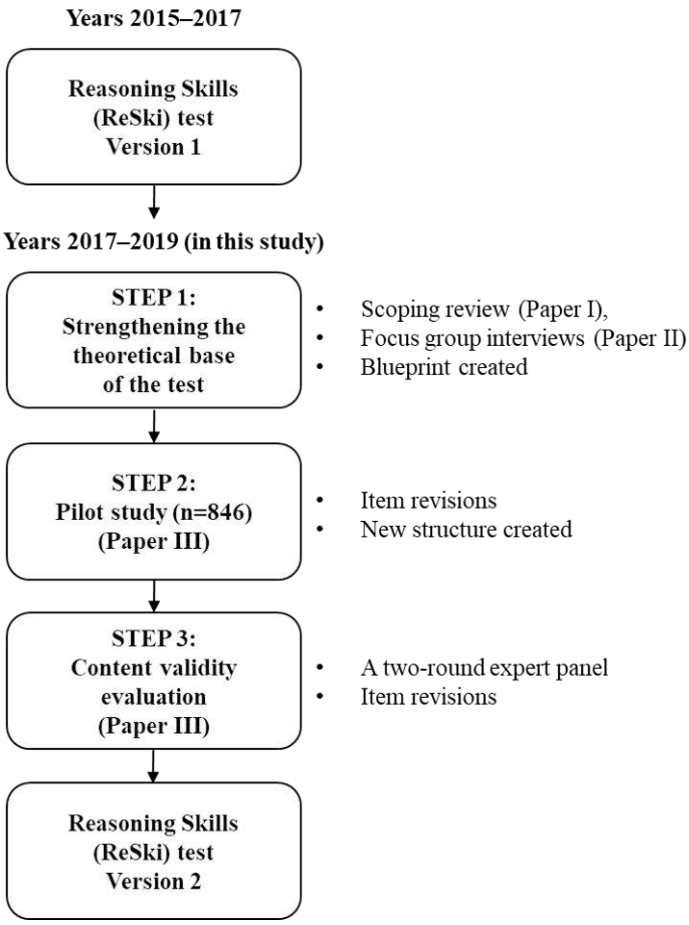

Figure 2. The development process of the ReSki test (version 2) (Modified from Paper III, Figure 1).

Reasoning Skills (ReSki) test

The first version of the ReSki test was developed prior this study (Figure 3). The second version of the test was developed in this study (Figure 2). Overall, the ReSki test was developed during the years 2015–2019 following the main principles of test development procedures (AERA, APA & NCME, 2014; DeVellis, 2017). The development process of the ReSki test has been described in detail in the Paper III.

The first version of the ReSki test was developed in the ReSSNE project (2015–2017) by the researchers to assess undergraduate nursing applicants' reasoning skills (Paper III). The test was developed for a joint electronic entrance examination that consisted of the domains of learning skills (including reasoning skills, language skills and mathematical skills), emotional intelligence and certainty of career choice, and thus the time limit and scoring of the ReSki test was allocated as part of the joint entrance examination. (Haavisto et al., 2019.) The test was initially developed based on previous literature (Levett-Jones et al., 2010; Simmons, 2010) and a synthesis of the results of a literature review and focus group interviews (Haavisto et al., 2019). The ReSki test was based on a reasoning process (Levett-Jones et al., 2010) and included a case followed by three question sections each including one closed-ended question related to the case with several response options (correct and distractor items) (Figure 3). The topic of the case was about overweight which was considered as a generic phenomenon without the requirement of previous nursing knowledge. The first version of the ReSki test was developed in two steps (Paper III, Figure 1).

The second version of the ReSki test was developed in this study during the years 2017–2019. Test development is an ongoing process of developing items and evaluating the psychometric properties of the test for which there is no endpoint (DeVellis, 2017; DeVon et al., 2007). Moreover, test version 1 (Figure 3) was originally developed simultaneously with all the domains of the joint entrance examination and there was a need to strengthen the theoretical base of the ReSki test. More detailed information about reasoning skills for the operationalisation of the concept and to ensure the development of a valid test was needed. The second version of the ReSki test was developed in three steps (Figure 2; Paper III, Figure 1).

In the first step, descriptive data were collected with a scoping review (Paper I) and with focus group interviews (Paper II) to identify reasoning skills to be assessed in undergraduate nursing student selection (Figure 2). Next, a blueprint was created to critically appraise the test (version 1) and to identify what is actually measured, and thus to support the development of the test (Figure 2; Paper III). The blueprint supported renaming the question sections to establish what is measured (Figure 3). More specifically, in strengthening the theoretical base of the test, the steps of the reasoning process were possible to identify to the selection phase (Paper II). In the focus group interviews, the beginning (steps 1–5) of the reasoning process (Levett-Jones et al., 2010; Paper II) was emphasised and therefore only the most relevant

and technically executable steps of 1) Collect information, 2) Process information and 3) Identify the problem and establish goals were included in the ReSki test (Paper III). More specifically, the step of consider the situation (Paper II) was not considered as relevant to be included to the test. According to previous literature, a similar decision has been made when structuring a simulation game of clinical reasoning by Koivisto et al. (2016). In addition, the steps of identifying the problem and establishing goals were both included to the question section three to maintain the assumed difficulty level of the test. Overall, the development process required constant consideration of the possible test difficulty, applicant perspective, features of the electronic platform and the use of the test as part of the joint entrance examination meaning the allocation of the test scores and timing as part of the overall exam scores and time.

In the second step, the ReSki test was piloted to analyse the level of discrimination and the difficulty of the items. After the pilot test, revisions were made for the structure and items of the test (refer to the data analysis, page 51) (Paper III, Figure 1).

In the third step, a two-round expert panel was conducted to evaluate the content validity. Based on the results, necessary revisions were made based on the panel results (refer to the data analysis, page 51) (Paper III, Figure 1).

As a result of the test development process, a ReSki test (version 2) assessing undergraduate nursing applicants reasoning skills according to reasoning process (Levett-Jones, 2010; Paper II) in nursing student selection context was developed (Figure 3). ReSki test is an electronic case-format test. Based on the case, a nursing applicant collects information, processes that information and finally makes a decision by identifying the problem and establishing goals. ReSki test includes three question sections each including one closed-ended question with 12 response options. Out of these 12 options, three items are correct and nine are distractor items (i.e. incorrect options that may be tempting to choose but are not correct). The test-taker is instructed about the number of correct items (e.g. choose three) and the scoring system. The ReSki test is performed under supervision. The time-limit and scoring of the ReSki test are part of the joint entrance examination which takes 2.5 hours. No penalty scores are given to avoid applicants' risk-taking strategies that may prevent measuring true ability because it has been found that for instance, high-achieving applicants with low degree of risk-taking behaviour may score lower than they usually would (Stenlund et al., 2018).

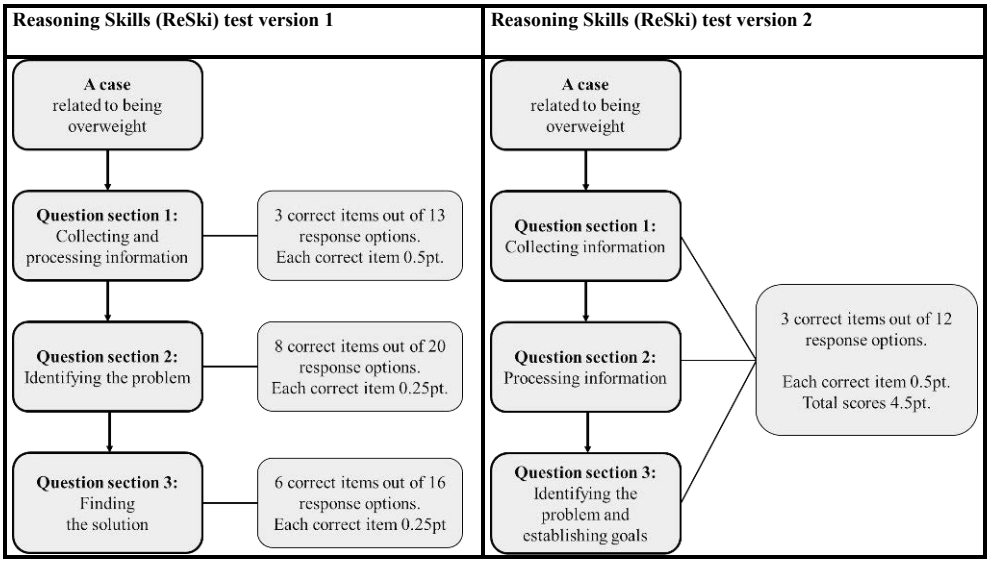

Figure 3. ReSki test versions 1 and 2.

Positive System Usability Scale (P-SUS)

A Finnish version of the P-SUS was used in this study to assess the perceived usability of the ReSki test (Table 1). The System Usability Scale (SUS) is widely used, since it was developed for a quick, reliable, low-cost usability scale that can be used for global assessments of how end-users perceive the systems usability (Brooke, 1996). Like the original SUS (Brooke, 1996), P-SUS is a quick, open access tool consisting of 10 items with five-point Likert scale (1=strongly disagree, 5=strongly agree) (Jokela, 2013). Like SUS, P-SUS provides a single overall SUS score ranging from 0 to 100 where 0 refers to the worst usability and 100 to the best usability (Brooke, 2013). The SUS score of 68 refers to an average usability (Brooke, 2013). The P-SUS is identical with the original SUS, but all the items are positively worded (Sauro & Lewis, 2011). The advantage of the positively worded SUS is the possibility to avoid mistakes in responding (i.e. the respondents may accidentally agree with the negative items) and to avoid miscoding (i.e. the researchers may forget to reverse the scales) (Sauro & Lewis, 2011). In this study, the Finnish version of P-SUS was used, and the main concept was modified referring to an electronic entrance examination system (Jokela, 2013).

4.4 Data analysis

Phase I

In the scoping review (Paper I), the data (n=24) were analysed by charting, collating and summarising (Armstrong et al., 2011) and using inductive content analysis (Elo & Kyngäs, 2008). First, characteristics of the selected studies (such as author(s), year of the publication and methodology), contents assessed and methods used to assess nursing applicant' learning skills, were charted and collated. Second, summarising and inductive content analysis were used to synthesise the data and answer to the RQs (Paper I). The analysis was conducted by the same researchers who performed the literature search and the final analysis was discussed in the research group.

In the qualitative descriptive study, the data from the focus group interviews (Paper II) were analysed with deductive and inductive content analysis (Elo & Kyngäs, 2008). The main categories were formed deductively and the subcategories inductively. The Clinical Reasoning Model by Levett-Jones et al. (2010) was used as a deductive framework to identify the main categories. In the deductive approach, the structure of analysis is operationalised based on previous knowledge (Elo & Kyngäs, 2008). The deductive approach with the Clinical Reasoning Model was chosen because the Clinical Reasoning Model describes the reasoning process step by step in nursing creating a detailed structure for the analysis. The Clinical Reasoning Model is a well-known framework and educational model that has been used in earlier studies (Levett-Jones et al., 2010; Theobald & Ramsbotham, 2019). The eight main steps (1) Consider the situation, 2) Collect cues and information, 3) Process information, 4) Identify the problem, 5) Establish goals, 6) Take action, 7) Evaluate outcomes and 8) Reflect on the process of action and new learning) of the clinical reasoning process were used as the main categories and thus the structure of the categorisation matrix. In this study, the nursing applicants' reasoning skills were analysed as generic skills not nursing specific. The data analysis was performed manually in a Word document. All in all, the data comprised 66 pages of transcribed interviews and 137 original phrases that answered to the research questions and were analysed. The original phrases from the interview data were collected to the categorisation matrix under the deductive categories for further condensing and coding. The codes were further synthesised into subcategories inductively. The main researcher conducted the analysis and the entire research group critically appraised the final analysis making changes with consensus.

Phase II

In the pilot study (Paper III), the data were analysed with descriptive statistics and with IRT (2PL) approach (Sulis & Toland, 2017). The data were analysed using Stata 15.1 (StataCorp., 2017) and Mplus 8.1 (Muthén & Muthén, 1998–2012) statistical programmes. Descriptive statistics (frequencies, percentages) were used to describe the item proportions in each question section. IRT analysis were used to analyse the discrimination and difficulty estimates (2PL) of the ReSki test (version 1) items. Correct items with >95% correct responses and distractor items with >99% correct responses were deleted from the analysis. The remaining items (43 out of 49) proceeded to the IRT analysis. As a result of the analysis, revisions were made for the ReSki test. Items with the difficulty estimate below –2 referring to very easy were deleted (Baker, 2001; Hambleton et al., 1991). Moreover, the structure of the ReSki test was modified to three correct items out of 12 options for each question section to unify the structure of the test and to reduce the possibility of guessing. As a result, revisions for the test were made since only 10 original items were accepted for further use (Paper III; Figure 3).

In the two-round expert panel evaluations (Paper III), the data were analysed using item-level Content Validity Index (I-CVI) (Polit & Beck, 2006) by summing up the experts' ratings. Percentages were used to describe the I-CVI and the acceptance limit was set for 80% in the first round and 100% in the second round (Imle & Atwood, 1988). More specifically, items/questions with less than 80% consensus were modified after the first round. In the second round, the experts agreed the clarity and relevance of the evaluated items/questions, resulting in 100% consensus and thus indicating an acceptable content validity for the ReSki test (version 2).

In the psychometrical testing of the ReSki test (Paper III), the data were analysed with descriptive statistics, correlation coefficients for subtotal/total scores, IRT modelling with discrimination, difficulty and pseudoguessing parameters and SUS-scoring technique for usability evaluation. The data were analysed using Statistical Analysis Software (SAS 9.4®) (SAS Institute Inc., 2015), Mplus 8.1. (Muthén & Muthén, 1998–2012), TestGardener online software and Microsoft Excel (Microsoft 365 ProPlus, version 2002). Participant demographics and ReSki test results were analysed with descriptive statistics (frequencies, percentages, mean, SD, range). The relationships amongst subtotals (the three question sections), relationships between the ReSki test total scores and the domain of learning skills (scores) and relationships between the ReSki test total scores and total entrance examination scores were analysed with Pearson correlation coefficient. The ReSki test items were analysed using IRT to examine the items' discrimination, difficulty and pseudoguessing estimates. Items (12 out of 36) with >95% correct responses were deleted from the analysis. First, the ReSki test items were analysed using 2PL IRT model showing

the difficulty and discrimination estimates. Second, as suggested by AERA, APA, & NCME (2014), a further analysis was undertaken to gather data to establish the validity of the test. The further analysis of the binary data was undertaken using TestGardener online software, a new version of the IRT (Li et al., 2019), to obtain more detailed ICCs and the three parameters of discrimination, difficulty and pseudoguessing (3PL). The 2PL model showed the discrimination and difficulty estimates for each item in all the question sections. The discrimination estimates (2PL) were classified according to Baker (2001) from very high to very low and 0 if not discriminating at all (Paper III, Table 5). The difficulty estimates (2PL) were classified according to Baker (2001) and Hambleton et al. (1991) from very hard to very easy. The 3PL model (TestGardener) showed the discrimination, difficulty and pseudoguessing parameters with more detailed ICCs for the ReSki test items. The difficulty estimates were recomputed and classified according to Ramsay et al. (2019) from easy to difficult based on the 5%, 25%, 50%, 75% and 95% quantiles. Pseudoguessing was considered high if it exceeded the 30% threshold (Ramsay et al., 2019; Tavakol et al., 2014). The pseudoguessing was analysed by observing the cut points in the 5% quantile in the ICCs, because there was a small amount of data available for estimating curve shapes below the bottom and over the top 5% intervals (Ramsay et al., 2019) (Paper III, Figure 2).

The perceived usability of the electronic ReSki test was analysed using the SUS-scoring technique (Brooke 1996; 2013). For the positively worded items, the score contribution is the scale position minus 1 meaning that each item's score contribution ranges between 0-4. Therefore, the P-SUS data were coded from the original range of 1–5 (Likert-scale) to the SUS-scoring technique range of 0–4. After calculating the individual item contributions, the sum scores were multiplied by 2.5 to obtain the single overall SUS score ranging from 0 to 100 (0=worst usability, 100=best usability) and a score 68 representing the average usability score (Brooke 1996; 2013). In addition, mean values for each re-coded P-SUS item were calculated. Only the questionnaires with full data (the participant answered all the ten items and there are no missing values) were included to the data analysis (Finstad, 2006).

Phase III

In the cross-sectional study (Paper IV), the data were analysed using Statistical Analysis Software (SAS 9.4®) (SAS Institute Inc., 2015). The demographic characteristics and results of the ReSki test were analysed with descriptive statistics (frequencies, percentages, measures of central tendency). Nursing applicants' reasoning skills were assessed with ReSki test score mean values on the level of above or below the centre of the range of possible scores, because the mean scores close to the centre of the range of possible scores are considered desirable (DeVellis,

2017). Pearson correlation coefficient was computed to assess the relationships amongst subscores. The factors related to nursing applicants' reasoning skills were analysed using analysis of covariance with Tukey's test in post-hoc multiple group comparisons (level of statistical significance at 0.05).

4.5 Ethical considerations

The study was conducted by following responsible conduct of research (The Finnish Advisory Board on Research Integrity TENK, 2012) including the protection of the data integrity (The European Parliament and the Council of the European Union, 2016a). The premises for the responsible conduct of research were followed in all the three study phases.

The study topic and its relevance to nursing science includes characteristics that needs to be considered from an ethical perspective. Nursing student selection affects thousands of HEIs and applicants on an annual basis. Assessment in student selection context is part of high-stakes testing because the selection decisions have consequences both for the applicant and the HEI. Therefore, the phenomenon of nursing student selection includes the elements of justice and equality. HEIs are responsible for using valid and reliable selection methods that enable the equal treatment and objective assessment of the applicants. This study has its basis in the ReSSNE project that was established nationally to improve nursing student selection practices. From an ethical perspective, the topic and aim of this study can be considered acceptable, because 1) there is a need to develop valid nursing student selection practices that are aiming for equal treatment of the applicants, 2) there is a need to select applicants who are most suitable for the education, proceed in their studies, graduate on time and deliver safe care for service users, and 3) this study produced ethically sustainably new knowledge and a novel test that can be used in further research in nursing science.

Phase I

In the first phase of this study, the scoping review (Paper I) did not require an ethical approval or permission, since the data included previously published studies that were publicly available via the databases. Honesty was aimed for in reporting the results and limitations of the scoping review.

In the qualitative descriptive study applying focus group interviews (Paper II), an ethical approval was provided by the Ethics Committee of the Higher Education Institution (11 September 2017). Permission to conduct the study was obtained from all the organisations involved and informed consent was obtained from all the study informants. The study informants received an information letter prior the interviews

and were again provided with the information in the beginning of the interviews. The study maintained voluntary participation, anonymity, and confidentiality of the study informants. Some of the expert informants knew the researcher beforehand, but there was no conflict of interest between the researcher and the informants. None of the student informants knew the researcher beforehand and were not researchers' own students. It was confirmed that the informants understood their rights and the meaning of their contribution to the study by asking the written informed consent. None of the informants withdrew from the study. Anonymity of the informants was ensured by storing the data carefully (i.e. only the researcher had the access), deleting personal, organisational, or otherwise identifying details from the transcripts and from the study report. The experts and the description concerning their organisations and positions were not reported in detail to ensure the anonymity of the informants. The data will be destroyed after the completion of the dissertational study.

Phase II

In the second phase of this study, the pilot study (Paper III) involved six Finnish UASs with 846 undergraduate nursing applicants. An ethical approval was provided by the Ethics Committee of the Higher Education Institution (20 August 2018). Permission to conduct the study was obtained from the participant UASs who were partners of the ReSSNE project. The applicants were informed about the study prior the data collection. The applicants received the invitation letter of the study together with the entrance examination invitation letter. The study maintained voluntary participation, anonymity, and confidentiality of the participants. Informed consent was obtained from the applicants electronically before they started the exam. The information about the applicants who consented/did not consent to the study was not provided to the UASs. The data were provided to the researcher via the project manager of the ReSSNE project who received the data (protected with a password) from the digital entrance examination system. The data were anonymised before the data analysis.

In the expert panel evaluations (Paper III), the permission to conduct the study was obtained from the experts' organisation. The same experts participated to the focus group interviews earlier in this study (the qualitative descriptive study in the phase I). The participants were informed about the study with a written study information letter. The study maintained voluntary participation, anonymity, and confidentiality of the participants. The experts consented to the study by answering to the e-mail.

In the psychometric testing of the ReSki test, ethical considerations, and procedures similar to pilot testing earlier in the phase II were followed. An ethical approval was obtained from the Human Sciences Ethics Committee in the Satakunta

region (former Ethics Committee of the Higher Education Institution) (14 April 2019), permission to conduct the study from the six UASs (partners of the ReSSNE project) and informed consent from the participating nursing applicants (n=1,056). The applicants received an invitation letter of the study together with the entrance examination invitation letter. Informed consent was obtained from the applicants electronically before they started the exam to ensure that the applicants understood their rights (voluntary participation, anonymity, and confidentiality of the participants) and the meaning of their contribution to the study. The participating UASs did not receive the information if their applicants consented or did not consent to the study. The data were provided to the researcher via the project manager of the ReSSNE project who received the data (protected with a password) from the digital entrance examination system. The data were pseudonymised before the data analysis and the original data including any identification details were stored behind the password with limited access.

Phase III

In the third phase of this study (Paper IV), the sample (1,056 nursing applicants) and the time of the data collection (Spring 2019) were the same as described in the psychometric testing of the ReSki test in the second phase of this study (Paper III). Therefore, the ethical considerations and procedures followed are the same as reported in the phase II (refer to the description of the psychometric testing in study phase II, pages 54–55).

5 Results

The results of this study are presented in this chapter according to the study phases and following the RQs 1–6. First, the results of the descriptive study phase are provided describing the reasoning skills to be assessed in nursing student selection (Paper I; Paper II; RQs 1–2). Second, the results of the test development phase are presented based on the psychometric testing and usability evaluation of the ReSki test (Paper III; Summary; RQs 3–4). Third, the results of the assessment phase are reported describing undergraduate nursing applicants' reasoning skills and factors related to them (Paper IV; RQs 5–6). Finally, a summary of the main results is provided. The more detailed descriptions of the results are presented in the original publications I–IV.

5.1 Reasoning skills to be assessed in nursing student selection

Reasoning skills assessed as part of learning skills in undergraduate nursing student selection (Paper I)

A scoping review with systematic literature search described the assessment of reasoning skills as part of learning skills in undergraduate nursing student selection (Paper I). The scoping review included 24 empirical studies (doctoral dissertations and research articles) published between 2006 and 2016. The chosen studies originated from four countries, mainly from USA (18 studies).

Four categories (language and communication, reasoning, mathematical and natural sciences skills) with several objects of assessment were identified for assessing nursing applicants learning skills (Paper I, Table 1). Language and communication skills, and mathematical skills were the most assessed learning skills whereas reasoning skills and natural sciences skills were assessed less. More specifically, seven studies reported the assessment of reasoning skills. Nine objects of assessment (analysis, inference, evaluation, critical thinking, decision-making, deductive reasoning, inductive reasoning, logic and problem-solving) were identified in reasoning skills (Paper I, Table 1).

The learning skills identified were assessed using two main methods: on-site selection methods (test/examination conducted before or during the selection process) and previous academic achievement (GPA of secondary school or pre-nursing studies) (Paper I, Table 2). On-site selection methods used included standardised tests, nationwide entry exam and interviews. Standardised tests were used most to assess applicants' learning skills. Reasoning skills were assessed using standardised tests (HSRT, WGCTA), nationwide entry exam (in Italy) and with interviews (MMI and an interview reported by Macduff et al., 2016).

Relationships between the on-site selection methods (most often standardised tests) assessing nursing applicants' learning skills and academic performance were reported in 19 of the chosen studies (Paper I, Table 2). The applicants' success in on-site selection methods were most often (reported in 15 studies) positively related to academic performance (i.e. success in the first semester/year or during the studies, attrition, timely graduation and passing the NCLEX-RN). In reasoning skills, nursing applicants' HSRT and WGCTA scores were positively related to academic success (first year and during the studies) and timely graduation whereas MMI scores were positively related to academic success during studies.

Overall, the results indicated that applicants' entry scores of learning skills were positively related to academic performance, but none of the existing selection methods assessed all four categories of learning skills. Overall scores in on-site assessments of learning skills (assessing more than only category) were the best predictor of future academic performance. Furthermore, reasoning skills in the selection phase were a promising predictor of future academic performance supporting the assessment of reasoning skills as part of other learning skills. However, the review results indicated that there is a need to clarify the concept of reasoning and its synonyms for the better operationalisation of the concept. In the chosen studies, the objects of assessment in reasoning skills were described on a general level lacking a more detailed description that could be used in the test development in the second phase of this study.

Reasoning skills identified for the selection of undergraduate nursing students (Paper II)

A qualitative descriptive study with focus group interviews (Paper II) identified the reasoning skills for the selection of undergraduate nursing students. The student informants of the study (n=16) were 21–50 years old, and mostly female (93.8%), whereas the expert informants (n=9) were 35–58 years old, mostly female (88.9%), and each had over 10 years of experience in healthcare (Paper II).

All eight steps of the clinical reasoning process (*Consider the situation, Collect cues and information, Process information, Identify the problem, Establish goals,*

Take action, Evaluate outcomes and *Reflect on the process of action and new learning*) (Levett-Jones et al., 2010) with 15 subcategories were identified as relevant for the student selection phase (Paper II, Figure 1). Although all the steps of the clinical reasoning process (Levett-Jones et al., 2010) were identified relevant for the selection phase, the first five steps were emphasised based on the number of original phrases analysed (Table 2). More specifically, information processing skills (Step 3) were most discussed and thus seen important to be assessed in nursing student selection. (Paper II; Table 2.)

Table 2. Reasoning skills identified for the selection of undergraduate nursing students.

Main categories*	**Subcategories**
Step 1: Consider the situation	Defining the situation Considering the whole situation
Step 2: Collect cues and information	Gathering information Gathering cues
Step 3: Process information	Interpreting the information Discriminating the information Synthesising the information Making inferences
Step 4: Identify the problem	Defining the problem
Step 5: Establish goals	Planning actions Finding the solution
Step 6: Take action	Implementing the solution
Step 7: Evaluate outcomes	Evaluating the decision
Step 8: Reflect on the process of action and new learning	Reflecting on one's actions Processing the feedback

* Most emphasised main categories are denoted with bolded text.

5.2 Psychometric properties and usability evaluation of the ReSki test

Psychometric properties of the ReSki test (Paper III)

The psychometric properties of the ReSki test (version 2) were tested with 1,056 undergraduate nursing applicants (Paper III). The majority (86%) of the applicants were female and slightly over half of the applicants (54%) were high school graduates (Table 3). Most of the applicants (59.5%) were first-time applicants and slightly over half of the applicants (51.5%) had nursing as their primary study

programme choice. Nearly half of the applicants (49.6%) had work experience in the field prior to application. (Table 3.)

Table 3. Characteristics of the participants (1,056 undergraduate nursing applicants) in study phases II and III.

Characteristics/Variable	n*	%
Age in years	1,050	-
Range	18–55	
Mean (SD)	24.56 (7.22)	
Gender		
Male	147	14.0
Female	904	86.0
Previous education		
High school	568	54.0
Vocational school	484	46.0
Previous Finnish higher education degree		
Yes	93	8.9
No	953	91.1
Previous application to nursing studies		
Yes	426	40.5
No	625	59.5
Study programme choice		
Primary choice	485	51.5
Not a primary choice	457	48.5
Work experience in the field prior to application		
Yes	520	49.6
No	529	50.4

* Missing values: Age in years (n=6), gender (n=5), previous education (n=4), previous Finnish higher education degree (n=10), previous application to nursing studies (n=5), study programme choice (n=114), work experience in the field prior to application (n=7).

The ReSki test demonstrated variance between the test-takers according to the mean values and SDs calculated for the sub-and total scores (Tables 4–7), which is considered important for high-stakes tests that need to be able to discriminate between the test-takers (DeVellis, 2017; Ramsay et al., 2020). In addition, the percentages of correct answers for ReSki test items (distractors and correct items) varied from 8.4% to 95%. The mean value of the ReSki total scores (2.72/4.5pt) was

slightly higher than the centre of the range suggesting that the overall difficulty of the ReSki test was acceptable, since a mean value close to the centre of the range of possible scores is considered desirable (DeVellis, 2017; Table 4).

Table 4. ReSki test total scores and number and proportion of correct answers as total.

Reasoning skills*	%	Range	Mean (SD)
Total scores	-	**0–4.5pt****	**2.72pt (0.80)**
Number of correct answers	-	**0–9**	**5.45 (1.59)**
Proportion of correct answers	**60.5**	-	-

* There were altogether nine correct items in the ReSki test. Each correct item was worth 0.5pt. Maximum total scores were 4.5pt.
** pt=points.

The question section one was rather easy for the applicants supported by the relatively high mean scores, and number and proportion of correct answers (Table 5).

Table 5. ReSki test question section 1: descriptive statistics of the responses for the distractor items and correct items.

Collecting information (Question section 1)	**f (of the correct responses)***	**% (of the correct responses)***	**Range**	**Mean (SD)**
Item 1	685	64.9	-	-
Item 2	**733**	**69.4**	**0–0.5pt****	**0.35pt (0.23)**
Item 3	>95% of the applicants got the item correct***			
Item 4	**953**	**90.3**	**0–0.5pt**	**0.45pt (0.15)**
Item 5	>95% of the applicants got the item correct***			
Item 6	>95% of the applicants got the item correct***			
Item 7	**463**	**43.8**	**0–0.5pt**	**0.22pt (0.25)**
Item 8	933	88.4	-	-
Item 9	984	93.2	-	-
Item 10	>95% of the applicants got the item correct***			
Item 11	>95% of the applicants got the item correct***			
Item 12	780	73.9	-	-
Subscores	-	**-**	**0–1.5pt**	**1.02pt (0.37)**
Number of correct answers	-	**-**	**0–3**	**2.04 (0.75)**
Proportion of correct answers	-	**67.8**	**-**	**-**

* Frequencies and percentages of the correct responses are reported for the distractor items and correct items. Only the correct responses for correct items were scored in the test.
There were three correct items in the question section. Each correct item was worth 0.5pt. Maximum subscores in the question section were 1.5pt.
The correct items in each question section are denoted with bolded text.
** pt=points.
*** The item was deleted from the further analysis for being extremely easy.

The question section two was the easiest question section of the test supported by the highest mean scores, and number and proportion of correct answers (Table 6).

Table 6. ReSki test question section 2: descriptive statistics of the responses for the distractor items and correct items.

Processing information (Question section 2)	**f (of the correct responses)***	**% (of the correct responses)***	**Range**	**Mean (SD)**
Item 1	>95% of the applicants got the item correct**			
Item 2	801	75.9	-	-
Item 3	1,003	95.0	-	-
Item 4	>95% of the applicants got the item correct**			
Item 5	**670**	**63.5**	**0–0.5pt*****	**0.32pt (0.24)**
Item 6	**812**	**76.9**	**0–0.5pt**	**0.38pt (0.21)**
Item 7	>95% of the applicants got the item correct**			
Item 8	>95% of the applicants got the item correct**			
Item 9	**849**	**80.4**	**0–0.5pt**	**0.40pt (0.20)**
Item 10	671	63.5	-	-
Item 11	>95% of the applicants got the item correct**			
Item 12	>95% of the applicants got the item correct**			
Subscores	-	-	**0–1.5pt**	**1.10pt (0.37)**
Number of correct answers	-	-	**0–3**	**2.21 (0.74)**
Proportion of correct answers	-	**73.6**	-	-

* Frequencies and percentages of the correct responses are reported for the distractor items and correct items. Only the correct responses for correct items were scored in the test.
There were three correct items in the question section. Each correct item was worth 0.5pt. Maximum subscores in the question section were 1.5pt.
The correct items in each question section are denoted with bolded text.
** The item was deleted from the further analysis for being extremely easy.
*** pt=points.

The question section three was the most difficult question section of the test supported by the lowest mean scores, and number and proportion of correct answers (Table 7).

Table 7. ReSki test question section 3: descriptive statistics of the responses for the distractor items and correct items.

Identifying the problem and establishing goals (Question section 3)	**f (of the correct responses)***	**% (of the correct responses)***	**Range**	**Mean (SD)**
Item 1	>95% of the applicants got the item correct**			
Item 2	515	48.8	-	-
Item 3	737	69.9	-	-
Item 4	927	87.9	-	-
Item 5	876	83.0	-	-
Item 6	**89**	**8.4**	**0–0.5pt*****	**0.04pt (0.14)**
Item 7	941	89.2	-	-
Item 8	**866**	82.1	**0–0.5pt**	**0.41pt (0.19)**
Item 9	875	82.9	-	-
Item 10	838	79.4	-	-
Item 11	866	82.1	-	-
Item 12	**315**	29.9	**0–0.5pt**	**0.15pt (0.23)**
Subscores	-	-	**0–1.5pt**	**0.60pt (0.35)**
Number of correct answers	-	-	**0–3**	**1.20 (0.70)**
Proportion of correct answers	-	**40.1**	-	-

* Frequencies and percentages of the correct responses are reported for the distractor items and correct items. Only the correct responses for correct items were scored in the test.
There were three correct items in the question section. Each correct item was worth 0.5pt. Maximum subscores in the question section were 1.5pt.
The correct items in each question section are denoted with bolded text.
** The item was deleted from the further analysis for being extremely easy.
*** pt=points.

All of the examined correlations indicated positive and statistically significant values (Paper III, Table 4). Correlations amongst the subtotals and between the subtotals and total scores suggested that if applicants scored high in one question section, they scored high in another ReSki question section and in the ReSki test as total, which supports the theoretical basis of the test. (Paper III, Table 4.) Correlations between the ReSki test total scores and the applicants' scores in the domain of learning skills (including reasoning skills, mathematical and language skills) ($r=0.44$, $p=<.0001$)

suggested that if applicants scored high in ReSki test, they scored high in the whole domain of learning skills, which supports the assumption of ReSki test measuring cognitive skills. Correlations between the ReSki test total scores and the applicants' total scores in the joint entrance examination (r=0.37, p=<.0001) indicated that if applicants scored high in ReSki test, they scored high in the whole entrance examination, which suggests that ReSki test identified high-achieving applicants, which is one important feature of an admission test (Ramsay et al., 2020).

The IRT analysis provided item-level information about the discrimination, difficulty and pseudoguessing of the ReSki test items (Paper III). First, the 2PL model provided the discrimination and difficulty estimates for the distractor items and correct items (Paper III, Table 5). The discrimination estimates varied from very high to very low. The discrimination values for the distractor items were low in comparison to the correct items. Two distractor items were highly discriminative, two were moderate and otherwise the distractor items had low or very low discrimination estimates. The discrimination estimates for correct items mainly ranged between moderate and very high, since only two correct items had low/very low discrimination estimates. The difficulty estimates varied from very easy to very hard. Most of the distractor items were easy for the test-takers, since only one distractor item had a moderate difficulty estimate and thus was a functional distractor. The difficulty estimates for correct items ranged between very easy and very hard. Based on the difficulty estimates, the question section three was the most difficult one for the test-takers. (Paper III, Table 5.) Second, a further analysis using 3PL model (TestGardener), provided the discrimination, difficulty and pseudoguessing parameters. More detailed ICCs for the test items were displayed (Paper III). Similar to the results of the 2PL model, the results of the 3PL model indicated that the test was mainly easy for the test-takers, except for question section three (Paper III, Table 5). Approximately half of the distractor items were susceptible for guessing amongst weaker examinees whereas only one correct item exceeded the 30% threshold. This suggests that most of the correct items measured the ability of the test-takers. Overall, the IRT results (Paper III) indicated that the quality of the correct items was quite good and the ReSki test as total was able to discriminate between the applicants. The items that were easy were most often distractor items. In addition, the distractor items demonstrated lower discrimination value and higher possibility for guessing compared to the correct items. (Paper III.)

Usability of the ReSki test (Summary)

The perceived usability of the electronic ReSki test (version 2) was assessed by undergraduate nursing applicants (n=849 out of 1,056 nursing applicants, response rate 80.4%). The overall usability of the ReSki test was assessed as acceptable based

on the P-SUS mean score (72.17, SD 28.92, range 0–100) being over the average usability level of 68 (Brooke, 2013) (Table 8). The lowest means scores per item were in the item one ("I think that I would like to use the system frequently") whereas the highest mean scores per item were in the item four ("I think that I could use the system without the support of a technical person") (Table 8).

Table 8. The perceived usability of the electronic ReSki test assessed by undergraduate nursing applicants (n=849).

P-SUS* items	**Mean (SD)**
1) I think that I would like to use the system frequently	2.59 (1.12)
2) I found the system to be simple	2.85 (1.28)
3) I thought the system was easy to use	2.97 (1.33)
4) I think that I could use the system without the support of a technical person	3.23 (1.41)
5) I found the various functions in the system were well integrated	2.79 (1.26)
6) I thought there was a lot of consistency in the system	2.78 (1.24)
7) I would imagine that most people would learn to use the system very quickly	3.03 (1.31)
8) I found the system very intuitive (=it was very easy to understand how the system works)	2.79 (1.27)
9) I felt very confident using the system	2.74 (1.24)
10) I could use the system without having to learn anything new	3.09 (1.39)
The P-SUS score**	**72.17 (28.92)**

* P-SUS=Positive System Usability Scale (Finnish version by Jokela, 2013. English version by Sauro & Lewis, 2011) using four-point Likert scale: range from 0 (most negative response) to 4 (most positive response).
** P-SUS score of 68 represent the average usability score (Brooke 1996, 2013).

5.3 Nursing applicants' reasoning skills and factors related to them

Nursing applicants' reasoning skills (Paper IV)

Undergraduate nursing applicants' (refer to the description of the psychometric testing in study phase II, page 43) were assessed with the electronic ReSki test (version 2) (Paper IV).

Nursing applicants' (n=1,056) total reasoning skills were slightly above the centre of the range of possible total scores (M 2.72pt/4.5pt, SD 0.80) and the proportion of correct answers in the test was 60.5% (Paper IV, Table 1; Table 4). Nursing applicants succeeded better in collecting information and processing information than in identifying the problem and establishing goals (Figure 4). Specifically, applicants scored above the centre of the range of possible subscores in collecting information (M 1.02pt/1.5pt, SD 0.37) and processing information (M 1.10/1.5pt, SD 0.37) whereas they scored below the centre of the range of possible subscores in identifying the problem and establishing goals (M 0.60pt/1.5pt, SD 0.35) (Figure 4; Tables 5–7). In addition, nursing applicants' reasoning skills varied between the applicants, since the SDs indicated variance between the applicant ability (subscores SD 0.35–0.37, total scores SD 0.80). (Paper IV, Table 1; Tables 4–7.) The high-achieving applicants' reasoning skills seemed to follow the reasoning process, evidenced by the statistically significant correlations amongst the subscores. This suggests that if the nursing applicant was able to collect the information, the applicant was able to process it and finally make the decision by identifying the problem and establishing goals (Paper IV, Table 2).

Figure 4. Nursing applicants reasoning skills according to their mean scores (M) in the ReSki test.

Factors related to nursing applicants' reasoning skills (Paper IV)

Age, gender, and previous education (high school/vocational school) were statistically significantly related to nursing applicants' reasoning skills (Paper IV, Table 3; Table 9). Work experience in the field prior to application was statistically significantly related to applicants reasoning skills only in identifying the problem and establishing goals (p=0.005) (Table 9). Other background variables examined did not indicate statistically significant results (Paper IV, Table 3).

Older applicants had better reasoning skills than younger applicants and high school graduates had better reasoning skills than applicants with vocational diploma (Table 9). Gender was statistically significantly related to nursing applicants reasoning skills, except in processing information. Overall, male applicants had better reasoning skills than female applicants. The applicants having work experience in the field prior to application (49.6%, n=520) were older (M 26.22, SD 7.50) than the applicants without previous work experience (M 22.87, SD 6.47) (Paper IV, Table 4). This suggests that applicants having work experience seemed to be older than the other applicants, which may explain their success in identifying the problem and establishing goals (Paper IV). (Table 9.)

Table 9. Factors statistically significantly related to undergraduate nursing applicants' (n=1,056) reasoning skills (Modified from Paper IV, Table 3).

Background variables	Subscores in Collecting information*	Subscores in Processing information*	Subscores in Identifying the problem and establishing goals*	Total scores in Reasoning Skills*
	Difference between means (95% confidence interval), **p-value**/NS=not significant			
Gender Male vs female	0.11 (0.03-0.18), **0.004**	-0.01 (-0.08-0.07), NS	0.11 (0.04-0.18), **0.003**	0.21 (0.05-0.36), **0.008**
Previous education High school vs vocational school	0.11 (0.05-0.17), **0.001**	0.10 (0.04-0.16), **0.001**	0.13 (0.07-0.19), **<.001**	0.34 (0.21-0.47), **<.001**
Work experience in the field prior to application Yes vs no	-0.00 (-0.06-0.06), NS	0.02 (-0.04-0.08), NS	0.09 (0.03-0.15), **0.005**	0.11 (-0.02-0.24), NS
	Slope (standard error), p-value/NS=not significant			
Age (continuous variable)	0.01 (0.00), **0.001**	0.01 (0.01), **0.001**	0.01 (0.00), **<.001**	0.02 (0.00), **<.001**

* Subscores (max. 1.5pt) and total scores (max. 4.5pt) of the ReSki test as a dependent variable (Analysis of covariance with Tukey's test in post-hoc multiple group comparisons).

5.4 Summary of the main results

This study identified reasoning skills to be assessed in undergraduate nursing student selection. An electronic test, measuring nursing applicants' reasoning skills (ReSki test), was developed and psychometrically tested (including the usability evaluation) and undergraduate nursing applicants' reasoning skills and factors related to them were assessed.

Based on the results of the scoping review (Paper I), reasoning skills were assessed in the selection phase and they were found to be a promising predictor of future academic performance. A comprehensive assessment of learning skills (i.e. assessing a wide range of learning skills predicted the academic performance best) were suggested. Furthermore, there was a need to clarify the concept of reasoning and its synonyms for the better operationalisation of the concept, and thus the scoping review did not provide enough description of reasoning skills that could have been used in the test development. (Paper I.)

Based on the results of the qualitative descriptive study with focus group interviews (Paper II), all eight steps of the clinical reasoning process (Levett-Jones et al., 2010) were identified as relevant for the student selection phase, meaning that the steps of the reasoning process were possible to identify to the selection phase. However, the first five steps of the reasoning process were more emphasised. (Paper II.) The results of the qualitative descriptive study were used to develop the ReSki test (version 2). The test was developed (Paper III) based on the most relevant steps of the reasoning process (collecting information, processing information, and identifying the problem and establishing goals).

Based on the results of the methodological cross-sectional study with psychometric testing (Paper III), the ReSki test (version 2) was valid, objective assessment of undergraduate nursing applicants' reasoning skills. The ReSki test was able to discriminate applicants and thus to rank them for the selection decisions. The correlations supported the theoretical structure of the ReSki test, and the electronic test had an acceptable usability. However, the IRT analysis suggested further development and testing focusing on the revisions of the distractor items to achieve the desired level of the test difficulty (Paper III).

Based on the results of the cross-sectional study (Paper IV), nursing applicants' reasoning skills vary. Nursing applicants' total reasoning skills were slightly above the centre of the range of total scores. Overall, the applicants were able to collect and process information, but they were less able to identify the problem and establish goals. Age, gender, and previous education were statistically significantly related to nursing applicants' reasoning skills indicating that older applicants, male applicants, and high-school graduates had better reasoning skills compared to the other applicants. Work experience in the field prior to application was statistically significantly related to applicants reasoning skills only in identifying the problem and establishing goals.

The main results of the study are visualised and summarised in Figure 5.

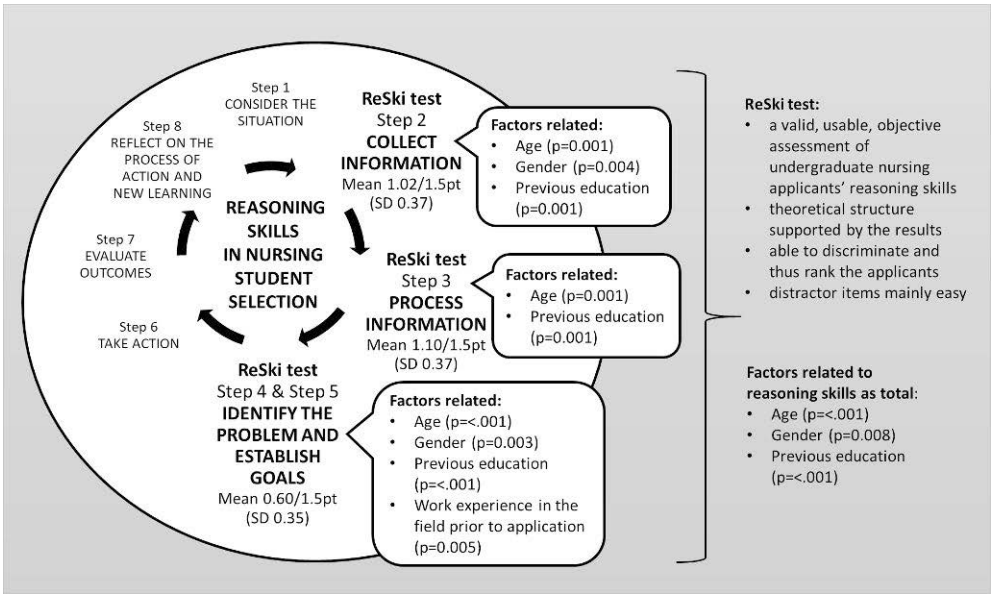

Figure 5. Summary figure of the main results (the steps of the reasoning process modified from Paper II, Figure 1; Levett-Jones et al., 2010).

6 Discussion

The results, validity and reliability of this study are discussed in this chapter according to the study phases and following the RQs 1–6. In addition, practical implications and suggestions for further research are provided. The more detailed discussions are presented in the original publications I–IV.

6.1 Discussion of the results

6.1.1 Reasoning skills to be assessed in nursing student selection

The purpose of the descriptive phase (phase I) of this study was to identify reasoning skills to be assessed in nursing student selection (Papers I, II). As a result of this study phase, reasoning skills for the student selection phase were identified. More specifically, the results of the scoping review (Paper I) indicated that reasoning skills have been assessed as part of other learning skills, but there is a need to clarify the concept of reasoning and its synonyms for the better operationalisation of the concept. The qualitative descriptive study with focus group interviews (Paper II) identified the reasoning skills for the selection of undergraduate nursing students according to the Clinical Reasoning Model by Levett-Jones et al. (2010).

The results of the scoping review (Paper I) confirmed the results of other studies: language, communication and mathematical skills are the most commonly assessed learning skills whereas reasoning skills together with natural sciences skills have been assessed less (Haavisto et al., 2019; Herrera, 2012; Wolkowitz & Kelley, 2010). Based on the results of the scoping review, learning skills are often assessed with standardised tests, which are most likely good predictors of future academic performance. However, the number of standardised tests in nursing student selection was overall rather small and standardised tests were mainly used in USA. It seems, that there is a gap in research on how standardised tests are used in nursing student selection in Europe. Furthermore, none of the standardised tests assessed all four categories of learning skills identified in this review. An important result of the scoping review was that the composite scores (i.e. the test/exam measured a range

of learning skills) predicted the academic success best, supporting previous literature in that student selection should focus on a variety of skills comprehensively reflecting the requirements of the professional education (Haavisto et al., 2019; Macduff et al., 2016; Schmidt & MacWilliams, 2011; Talman et al., 2018a; Taylor et al., 2014; Wambuguh et al., 2016). The results highlight the need to develop standardised and objective assessment methods focusing on a variety of nursing applicants' learning skills. Furthermore, reasoning skills were a promising predictor of future academic performance supporting the assessment of reasoning skills as part of a comprehensive assessment of learning skills in nursing student selection (Paper I). However, reasoning skills in nursing student selection were not identified on a detailed level and the concepts related to reasoning skills (e.g. decision-making, critical thinking, problem-solving) were often used as synonyms. This result is in accordance with previous studies (Georg et al., 2018; Haavisto et al., 2019; Hernandez, 2011; Lauri & Salanterä, 2002; Simmons, 2010). Based on the review results, the objects of assessment in critical thinking seemed to be identified best to nursing student selection (Paper I). Nevertheless, a variety of definitions of critical thinking exists proposing that critical thinking is a a facilitator of unbiased reasoning involving both cognitive skills and dispositions of the mind (Alfaro-LeFevre, 2013; Carbogim et al., 2016; Ennis 1985; Facione 1990; Facione et al., 1994; Heijltjes et al. 2015; Hong & Yu, 2017; Simmons, 2010). Therefore, it seems that critical thinking refers to both cognitive and non-cognitive skills/attributes. For this reason, the scoping review indicated the need to study reasoning skills as cognitive skills in student selection context further, because it may be challenging to assess dispositions of the mind objectively.

The results of the qualitative descriptive study with focus group interviews (Paper II) indicated that reasoning skills relevant to the student selection phase follow the clinical reasoning process (Levett-Jones et al., 2010). Overall, similar to clinical reasoning process, reasoning skills in the selection phase were a step-by-step process that especially involves collecting and processing information leading to decision-making by identifying the problem and establishing goals (Levett-Jones et al., 2010; Paper II). Based on the results, the focus of the assessment of nursing applicants' reasoning skills should be placed in the beginning of the reasoning process and particularly on information processing skills (Paper II). The final steps of the reasoning process (i.e. the steps of take action, evaluate outcomes and reflect on the process of action and new learning) (Levett-Jones et al., 2010) were emphasised less in the focus group interviews. A reason for this result may be that the informants did not consider these steps relevant to the selection phase. Another interesting result was that the study informants also recommended the collection/assessment of visual and auditory cues alongside written information which should also be considered in the future (Paper II). All in all, the results of the

qualitative descriptive study (Paper II) reflected the increasing cognitive requirements of health care environments where the importance of information processing skills is highlighted (FIOH, 2018; Levett-Jones et al., 2010; Simmons, 2010). The results of the qualitative descriptive study (Paper II) were used in the test development in the second phase of this study (Paper III).

6.1.2 Psychometric properties and usability evaluation of the ReSki test

The purpose of the test development phase (phase II) of this study was to develop and psychometrically test a ReSki test for undergraduate (bachelor level) nursing student selection (Paper III). As a result of this study phase, the electronic ReSki test (version 2) was developed (refer to the description of the test development process, pages 46–49) and tested for its psychometric properties and usability. Presumably, the ReSki test is a first test to assess nursing applicants' reasoning skills according to the reasoning process. Furthermore, the IRT approach, which has been recommended but rarely used in nursing science for the validity evaluation, was used as a main method to assess the psychometric properties of the ReSki test (Tavakol et al., 2014).

Based on the results of the psychometric testing, the ReSki test is a novel, valid objective assessment of undergraduate nursing applicants' reasoning skills (Paper III). The content validity of the ReSki test (assessed by the expert panel) was considered acceptable in the development process of the ReSki test and the further psychometric testing provided support for the reliability and validity of the ReSki test. First and foremost, the ReSki test demonstrated variance in applicants' results and the test was able to discriminate and thus set the applicants to rank order, which is a basic requirement for a valid and reliable admission test (Ramsay et al., 2020). Second, the results, namely, correlations supported the theoretical basis and structure of the test, indicating that reasoning skills are cognitive skills and decision-making is based on collecting and processing information (Levett-Jones et al., 2020, Simmons, 2010). Third, the IRT approach enabled to focus on the quality and characteristics of individual test items by providing information on the test difficulty, discrimination and pseudoguessing. The results of the IRT analysis supported the descriptive results of this study phase indicating that the ReSki test was relatively easy for the applicants. This may be because the distractor items were not tempting for most of the applicants to choose. According to previous literature, test-takers can be cleverer than the items and it is a challenge for test-developers to construct good items and thus functional distractors (DeVellis, 2017; Tavakol et al., 2014). However, the question section three (identifying the problem and establishing goals) was more difficult than the previous question sections, because the quality of the

distractor items were better. As a result, the test achieved a more satisfactory overall test difficulty (Paper III). Based on the results of the IRT analysis, the distractor items would need further revision to achieve a more desired level of difficulty for the ReSki test. Typically, an easy item has lower discrimination and higher possibility of guessing as well (Tavakol et al., 2014). Therefore, it is important to focus on the adjustment of the difficulty level of the ReSki test. In addition, the correct items in the question section three should be critically appraised, since the results indicated that one of the items could be even too difficult for the applicants (Paper III). The study results support the statement, that developing quality items is an iterative, time-consuming process of developing and testing the items (DeVellis, 2017).

Based on the results of the usability evaluation, the electronic ReSki test has an acceptable usability. Overall, the results indicated that an electronic test can be applicable in student selection. More specifically, the results indicated that taking the electronic test was fluent for the applicants, which is important when developing high-stakes tests aiming for fair and fluent student selection processes (Haavisto et al., 2019; Ministry of Education and Culture, 2016; Shulruf et al., 2018; Talman, 2014). The use of electronic tests in nursing student selection context has been rare, at least in Finland, and therefore it was important to evaluate usability together with the psychometric properties. In the future, electronic entrance examination tests could assist cost-effective selection practices as well. Furthermore, electronic tests could use automated scoring, which may reduce the errors observed in manually calculated exam scores.

6.1.3 Nursing applicants' reasoning skills and factors related to them

The purpose of the assessment phase (phase III) of this study was to assess nursing applicants' reasoning skills and factors related to them (Paper IV). Reasoning skills of the nursing applicants were successfully assessed by using the ReSki test (version 2).

Based on the results, nursing applicants total reasoning skills were rather good based on their total score results that were over the centre of the range of possible scores (Paper IV). This result was in a line of a previous study (Pitt et al., 2015) focusing on nursing students' critical thinking skills in which participants' total HSRT entry scores demonstrated a midrange ability (Facione et al., 2011). Nevertheless, the results of this study (Paper IV) indicated that nursing applicants reasoning skills varied, and the applicants were less able to identify the problem and establish goals than collecting and processing information. According to the results, the high-achieving applicants possessed reasoning skills according to the reasoning

process (Levett-Jones et al., 2010; Paper II) meaning that they collected and processed the information finally making the decision by identifying the problem and establishing goals. It is understandable that not all applicants were able to identify the problem and establish goals, because comprehensive thinking is needed in the reasoning process meaning that decisions are based on the information that is collected and processed (Levett-Jones et al., 2010; Paper II). This means that correct choices are needed in previous steps to later identify the problem and establish goals. A similar result was found in a previous study focusing on nursing students' learning of reasoning skills in clinical scenarios, which stated that students learned how to collect information but were less successful in learning to establish goals (Koivisto et al., 2016). In a previous study (Paper II) from Phase I of this doctoral dissertation, the results suggested that the focus of assessment of nursing applicants' reasoning skills should be on information processing skills. However, based on the results of the phase III (Paper IV) and supported by the results of the psychometric testing (Phase II), it seems important to assess reasoning skills according to the reasoning process rather than solely assessing a specific step of the process (such as information processing skills).

In this study, age, gender, and previous education were statistically significantly related to the reasoning skills of nursing applicants (Paper IV). Additionally, work experience in the field of nursing prior to application was statistically significantly related to reasoning skills, but only in identifying the problem and establishing goals. Specifically, the results indicated that older applicants, male applicants, and high-school graduates had better reasoning skills than younger applicants, females, and applicants with vocational diploma. These results are important, since it is important to understand the role of demographic factors related to the success in student selection. This may help HEIs in developing fair and evidence-based student selection practices. The results of this study are in accordance with previous studies as statistically significant relationships between entry scores and age and gender have been reported (Pitt et al., 2015; Stage & Ögren, 2004). For example, Pitt et al. (2015) reported that on entry to the pre-registration nursing programme older participants had significantly higher sub-scale evaluation scores. In SweSAT, both age and gender differences have been found being rather typical for standardised tests. However, it is possible to tackle the influence of demographic variables, such as age and gender, and provide a fair test for a heterogenous group of applicants by constructing entrance examinations that assess applicants' skills comprehensively and not focusing on a narrow field of ability. (Stage & Ögren, 2004.) Furthermore, an important result was that high-school graduates possessed better reasoning skills than applicants from vocational background both in every step of the reasoning process and in the ReSki test as total. This result indicates that high-school graduates may have a better basis for developing their reasoning skills in the professional

context and a better basis for entering the nursing programmes, although all secondary education institutions should prepare students for higher education studies in Finland and in other EU member states. The participants of this study with a vocational diploma and work experience in the field prior to application were most likely practical nurses. Furthermore, the result of work experience in the field prior to application being statistically significantly related to identifying the problem and establishing goals, could be explained by applicants' higher age rather than the work experience itself based on a descriptive analysis. Previously, Pitt et al. (2015) reported that students with prior nursing-related experience had significantly lower entry critical thinking scores, suggesting that assistant or helper roles did not develop higher-level thinking skills. It is possible that practical nurse education or work as a practical nurse, does not necessarily include complex decision-making where the use of a cognitive process is needed to make judgements about the patient care. In general, practical nurse education has been rarely studied in nursing education research and little is known about their education from evidence-based perspective (Vierula et al., 2016). According to previous literature, prior work experience in the field has been used as an admission criterion for higher education (Capponi & Mason Barber, 2020; Schmidt & MacWilliams, 2011). The results of this study do not support the use of this admission criterion and suggest a need for further research concerning the use of prior work experience as an admission criterion. All in all, the results indicate that nursing applicants' reasoning skills vary, which suggests that students entering the nursing programme may have different kinds of learning needs in the beginning of their studies.

6.2 Validity and reliability of the study

The validity and reliability of this study were considered throughout the study in all its phases. The aim of the study was to develop a valid and objective ReSki test for national use in nursing student selection and thus the study followed the main principles of test development procedures (AERA, APA & NCME, 2014; DeVellis, 2017). According to AERA, APA & NCME, (2014), validity is the most fundamental consideration in developing and evaluating tests. Accumulation of relevant evidence is recommended for a sound scientific basis of a high-stakes test and the use of the test scores. Well-constructed, valid tests benefit the test-takers and test-users (HEIs) supporting better selection decisions and a more equitable access to education. However, validity is a broad concept and refers to a wide interpretation of test scores and thus is more than simply "the validity of a test". (AERA, APA & NCME, 2014.)

In addition, reliability is a complex term in the test development context since it has either been used as a very general concept or is referring to reliability coefficients

of CTT. Furthermore, CTT is not recommended as the main approach to assess psychometric properties of tests, especially high-stakes tests (AERA, APA & NCME, 2014; Tavakol et al., 2014). In this study, reliability was taken into account by observing the variance of the items (i.e. response options) and test scores, and identifying high-quality items and thus the reliability of the test was considered as part of the test validity (AERA, APA & NCME, 2014; DeVellis, 2017; Devon et al., 2007; Ramsay et al., 2020). Several methods, approaches and data collections were used in this study to gather accumulated evidence of validity and reliability of the ReSki test.

Phase I

In the scoping review (Paper I), the issues concerning validity and limitations of the study were related to problem identification, literature search, data analysis and presentation of the results (Whittemore & Knafl, 2005). Overall, Armstrong et al.'s (2011) steps of conducting the scoping review were used to follow the research process systematically and to ensure valid results. In the problem identification (Whittemore & Knafl, 2005), the entire research group participated in the formulation of the RQs, search terms and the inclusion/exclusion criteria, which can be considered strengthening the validity of the research process and thus validity of the results (Armstrong et al., 2011; Aveyard, 2007). In addition, the RQs were clearly identified prior the data search. In the literature search (Whittemore & Knafl, 2005), several preliminary searchers were conducted, and various synonymous concepts were searched from dictionaries and other literature before establishing the final search terms. A thorough preliminary search process strengthened the validity of the search process. In the final search from several databases including manual search, two researchers (Paper I) conducted the identification, selection and charting of the data to ensure the inclusion of relevant studies and accuracy of the collected information. A limitation of the study relates to the literature search, because it was limited to articles published during the years 2006–2018 to ensure the inclusion of up-to-date studies. However, this may have excluded some relevant studies. Concerning the data analysis and presentation of the results, the analysis was conducted together with the research group to ensure a correct interpretation of the results of the studies included in the review. In addition, the results of the review for further research and practice were discussed enhancing the meaningfulness and trustworthiness of the results (Whittemore & Knafl, 2005). For the validity and generalisability of the review results, it should be noted that the selected articles reported the reliability and validity of the used methods/tests poorly and most of the studies originated in the USA. Additionally, previous research has reported the use of synonymous terms in studies focusing on student selection and/or reasoning

(Haavisto et al., 2019; Hernandez, 2011; Simmons, 2010), which should be considered when interpreting the results of this scoping review. Overall, the results can be considered as generalisable due to the more or less common requirements of higher education studies and nursing competencies.

The qualitative descriptive study with focus group interviews (Paper II), was evaluated for its trustworthiness (credibility, dependability, and transferability) of data collection, data analysis and interpretation of the findings (Doody et al., 2013c; (Graneheim & Lundman, 2004). The possible limitations of this study were related to the abstract study topic and objectiveness of the findings. In all the interview studies, the findings should be truthful representing the ideas of the informants and context of the discussion not a reflection of the researcher's biases (Doody et al., 2013c; Turner, 2010). Therefore, staying objective is a challenge of interview studies and the role of the moderator is important in focus group studies to collect rich and valid insights from the group of informants (Stewart et al, 2007). Concerning the data collection, the credibility of the study was established by choosing focus group method as the most suitable method for the data collection. The dependability was not threatened, since the data collection period was relatively short. (Graneheim & Lundman, 2004.) The focus group interviews were planned carefully beforehand together with the research group. To ensure rich focus group discussions, an interview guide was planned, piloted, and used (Doody et al., 2013b; Vaughn et al., 1996). Although the pilot interviews were informal in their nature and were not recorded, the researcher had a possibility to practice the role of the moderator, which can be considered important for the collection of rich and objective data (Doody et al., 2013b). Nevertheless, it is possible that the use of only one moderator has led to the risk of bias in objectivity. To avoid the subjectivity and to gather rich data, the moderator used reflective listening. In the interviews, it was obvious that the experts provided richer and more relevant (i.e. detailed and concrete descriptions) data than the graduating students. Nevertheless, group interaction enabled student informants to generate distinct data, which at the end, is the primary goal of focus group interviews (Doody et al., 2013a). The students provided relevant complementary insights to the expert opinions and enabled a comprehensive data (Doody et al., 2013b). However, based on the number of original phrases analysed and as reported in the results (refer to the chapter 5.1, page 58), the data were richer in the first five steps (i.e. main categories) than in the steps 6–8. The number of experts in the focus groups was smaller than the students, because it was more challenging to find expert informants than students. The trustworthiness of the data analysis was enhanced by conducting the analysis together with the research group. In addition, the analysis was initiated deductively according to a well-known framework (Clinical Reasoning Model by Levett-Jones et al., 2010) used in earlier studies (Theobald & Ramsbotham, 2019) and recommended as a systematic approach in educational

scenarios (Cook et al., 2010; Koivisto et al., 2016; Levett-Jones et al., 2010; Petit dit Dariel et al., 2013; Shipman et al., 2012). In general, deductive analysis has been quite rarely used in nursing education research in Finland (Vierula et al., 2016) and much less than inductive approach (Elo & Kyngäs, 2008). The strength of the deductive analysis in this study was that it enabled objectivity and operationalisation of an abstract concept on the basis of previous knowledge (Elo & Kyngäs, 2008). However, the subcategories were formed inductively and the awareness of the original descriptors in the Clinical Reasoning Model (Levett-Jones et al., 2010) may have limited the full conformability of the findings in the inductive approach (Paper II). This should be considered when interpreting the study findings. Overall, for the face validity of the results, the research group considered them credible. According to Plummer-D'Amato (2008), focus group results are difficult to be generalised to a larger population, because of the purposive sampling and small sample sizes. Nevertheless, the findings are considered transferable to nursing student selection context due to the relatively similar requirements of higher education studies and nursing competencies.

Phase II

In the test development phase (phase II) of this study (Paper III), both the development process of the ReSki test and the psychometric testing required a constant consideration of validity and reliability.

In developing the ReSki test (version 2) (refer to the description of the test development process, pages 46–49) rigorous steps (i.e. strengthening the theoretical base, pilot testing and content validity evaluation) were taken to develop the test, which can be considered to enhance the validity of the overall development process (AERA, APA & NCME, 2014). The theoretical base of the test was thoroughly studied because the operationalisation of the concept has been considered challenging in previous studies (Carbogim et al., 2016; Zuriguel Pérez et al., 2014). The pilot study included a large sample size for a pilot sample, and it enabled the 2PL IRT analysis and thus revisions for the test especially related to item difficulty. The content validity evaluations by the expert panels were used thoroughly during the development process of the ReSki test (versions one and two). The first expert panel was used prior this study in developing the first version of the test. In this study, a two-round expert panel was used (Polit & Beck, 2006). However, the expert panel was rather small and although following the methodological recommendations (Imle & Atwood, 1988; Polit & Beck, 2006), a panel with more experts may have provided more insights and objectivity for the development process. Overall, the expert panel supported the content validity of the ReSki test.

In the psychometric testing (Paper III), the issues concerning validity and reliability of the study were related to data collection, data analysis and interpretation of the results. The results supporting the ReSki test being a valid and reliable test including the suggestions for revisions have been reported in the Paper III (refer to the chapter 5.2, pages 58–64 and chapter 6.1.2, pages 73–74). In the psychometric testing, the response rate was 55.4%. Although almost half of the applicants did not participate to the study, the sample (n=1,056) can be considered representative to the study population. The sample size was adequate for IRT analysis, large in comparison with sample sizes in other similar nursing student selection studies (Paper I), and the sample represented typical characteristics of the population (Paper III). In addition, the six Finnish UASs (out of 23) that participated to the study were nationally widely represented. In the data collection, an electronic test was used enabling to minimize errors in data input. The usability of the electronic ReSki test was evaluated with P-SUS (Jokela, 2013), which has been evaluated as valid and reliable scale and applicable to several contexts of usability evaluation (Brooke 1996, 2013; Sauro & Lewis, 2011). Concerning the data analysis, IRT approach was used as the main method for the validity evaluation as recommended, that is a major strength in this study (Tavakol et al., 2014; Yang & Kao, 2014). In addition, several analysis methods (as recommended by the AERA, APA & NCME, 2014) were used to support the results of 2PL and 3PL IRT analysis and thus the validity and reliability of the ReSki test. In addition, the validity of the study results is supported by the fact that the analysis was conducted together with statisticians and by consulting experts of psychometrics (IRT analysis). For the interpretation of the results and their validity, concerning item discrimination, difficulty and pseudoguessing parameters, it should be noted that the interpretation of the results is depended mainly on the ICCs (e.g. DeVellis et al., 2017). Specifically, methodological literature does not provide strict guidelines on the ideal level of difficulty or the maximum of the pseudoguessing level in tests. In this study, an overall test difficulty level was considered acceptable when applicant mean performance was close to the centre of the possible score range and the item pseudoguessing level was considered rather high if exceeding the possibility of 30% threshold. (DeVellis et al., 2017; Ramsay et al., 2020; Tavakol et al., 2014.) The validity of an instrument could be evaluated from other perspectives as well. In this study, face validity of the ReSki test was not evaluated to avoid subjectiveness of the evaluation (DeVon et al., 2007). Criterion validity was not evaluated in this study, since a lack of a suitable parallel instrument and the context of student selection where applicants are participating to a high-stakes tests with a time-limitation (DeVon et al., 2007; Lakanmaa, 2012). This study did not evaluate predictive validity as longitudinal research design was not used (DeVon et al., 2007). However, predictive validity is an essential characteristic of a valid admission test to fulfil the

aims and requirements set for student selection and thus recommended to be studied in the further development of the ReSki test.

Phase III

In the assessment phase (phase III) assessing nursing applicants' reasoning skills and factors related to them (Paper IV), the issues of validity and reliability are consistent with the phase II of this study, since the ReSki test (version 2) was used to assess reasoning skills in the same study sample than in the phase II (refer to the chapter 6.2, page 79–81). Specifically, the ReSki test was suitable to assess reasoning skills of nursing applicants and the sample was considered representative. The sample size was adequate to conduct statistical analysis, and the relevant statistical methods were chosen together with statisticians. The studied background variables were based on previous studies (Paper IV) enhancing the validity of the results. The study results are generalisable nationally. Internationally, the study context (i.e. the assessment of reasoning skills) is generalisable, because of the more or less common requirements of higher education studies and nursing competencies. However, further research is needed concerning nursing applicants reasoning skills and relating factors in other than Finnish populations. Cross-sectional design limits drawing of conclusions about any strong causality of the associations between the variables studied. Nevertheless, the results indicate that some background factors are related to reasoning skills in the selection phase. The strength of this study was that the background variables and their relationships to applicants' reasoning skills were identified. Student selection methods should be valid and reliable, but also fair between the applicants requiring a need to study factors related to the assessed skills.

6.3 Practical implications

Based on the results of this study, the following practical implications for education can be presented:

- HEIs are encouraged to assess nursing applicants' reasoning skills as part of a comprehensive assessment of learning skills. The comprehensive assessment could assist HEIs to select nursing applicants who would be successful in their studies. Most likely, the comprehensive assessment of learning skills could also benefit other higher education disciplines than nursing, due to the similar requirements of higher education studies.
- HEIs are encouraged to assess nursing applicants' reasoning skills as generic, cognitive skills according to the reasoning process including collecting information, processing information, and identifying the

problem and establishing goals. The assessment of these steps of the reasoning process in nursing student selection does not require any professional knowledge. For this reason, the assessment of reasoning skills does not require long preparation from the applicants, and thus supports fair selection practices that fulfil educational policy requirements in Finland.

- HEIs are encouraged to use standardised tests for valid and objective assessment of nursing applicants and to ensure fair selection practices.
- The ReSki test can be used to assess undergraduate nursing applicants' reasoning skills. The use of valid and reliable tests in nursing student selection is important for the objective assessment and equal treatment of applicants. In addition, the selection of applicants most likely to succeed and proceed is beneficial for future nursing practice.
- Electronic entrance examination tests may be usable for test-takers and could possibly assist HEIs in developing fluent and cost-effective selection practices.
- It is encouraged to critically appraise how vocational education currently prepare students for higher education studies.
- This doctoral dissertation study focused on reasoning skills in nursing student selection. In the future, nurse educators could possibly use the ReSki test results of the applicants as a baseline information or conduct the ReSki test to first-semester nursing students when planning their first-year studies. HEIs should pay attention to the teaching and learning of reasoning skills as a process from the beginning of the degree to ensure nursing students' adequate reasoning skills in complex decision-making and clinical scenarios as graduating nurses.

6.4 Suggestions for further research

Based on the results of this study, the following suggestions for further research can be presented:

- Further research is needed to develop and test the ReSki test further. Revisions, especially concerning the distractor items, are suggested to achieve a more desired level of difficulty. Furthermore, the predictive validity of the ReSki test should be studied with a longitudinal design.
- The ReSki test could be used for wider research purposes in the future. It could be modified to other educational contexts in nursing as well, by shifting the focus from generic to nursing-specific assessment of

reasoning skills and thus following the Clinical Reasoning Model (Levett-Jones et al., 2010).

- New standardised tests and thus objective assessment methods measuring nursing applicants' reasoning skills and other learning skills should be developed. The tests could utilise more sophisticated electronic platforms. In the assessment of nursing applicants' reasoning skills, the tests could also comprise the collecting of other cues (such as auditory and visual) than written information, but taking into account the accessibility directive (The European Parliament and the Council of the European Union, 2016b).
- The evidence of using standardised tests in nursing student selection is biased to USA. Member states of EU and other countries are encouraged to provide research on objective assessment methods in their populations as well.
- This study operationalised and identified reasoning skills for the student selection phase, but also confirmed the complexity and ambiguous use of the synonymously used concepts of reasoning. Further research is recommended concerning reasoning and synonymously used terms for the systematic use of these concepts. For example, the Clinical Reasoning Model (Levett-Jones et al., 2010) is suitable to be used as a deductive framework in research. The model could assist researchers to identify reasoning skills both as generic and professional skills in various scenarios.
- IRT approach is recommended to be used in the future studies focusing on the development and psychometric properties of high-stakes tests. Furthermore, a greater understanding of the IRT methods could benefit researchers and educators to develop valid tests and identify high-quality items.
- All the learning skills, not only reasoning, and methods assessing them, should be studied with longitudinal designs in relation to academic performance and especially to clinical success, since the predictor variables are often theoretical or general in their nature (GPA, attrition, graduation etc.).
- This study indicated that nursing applicants reasoning skills vary in the selection phase and thus further research could focus on how reasoning skills develop during nursing education, and how they can be assessed and supported.

- Further studies could also focus on vocational education and seek to identify the possible gaps in preparation to higher education studies.

7 Conclusions

This study produced a ReSki test (version 2) and new knowledge concerning undergraduate nursing applicants' reasoning skills and the assessment of these skills in nursing student selection.

As a conclusion, this study suggests that nursing applicants' reasoning skills are cognitive, generic skills and recommended to be assessed according to the reasoning process involving the steps of collecting information, processing the collected information, and identifying the problem and establishing goals. The assessment of reasoning skills as part of a comprehensive assessment of learning skills and assessed with standardised tests, may assist HEIs to develop evidence-based, objective, and fair student selection practices. The developed ReSki test is a valid, usable, and objective assessment of undergraduate nursing applicants' reasoning skills. The IRT approach was successfully used in this study to assess the psychometrics of the ReSki test and the IRT analysis provided valuable item-level information. These results can be used to develop the ReSki test further, since there is a need for the revisions of the distractor items in the adjustment of a more desired difficulty level for the test. Moreover, nursing applicants' reasoning skills vary in the selection phase indicating that some nursing applicants may enter the nursing programme with better reasoning skills than other applicants. Nursing applicants are better in collecting and processing the information than in identifying the problem and establishing goals highlighting the importance of assessing and teaching of reasoning skills according to the reasoning process. Nursing applicants reasoning skills in the selection phase are related to some background variables, most importantly to previous education indicating that vocational education is not necessarily developing adequate reasoning skills and thus preparing students for higher education studies.

The study results have implications for nursing education and research and thus for nursing practice and education policy.

Acknowledgements

This study was carried out at the Department of Nursing Science, University of Turku, Finland. It truly took a group of remarkable people to inspire all of this. I would like to express my deepest gratitude to all those who made this doctoral dissertation study possible, although I cannot name everyone here individually.

First and foremost, I would like to express my deepest gratitude to my supervisors Professor Elina Haavisto, PhD Maija Hupli and PhD Kirsi Talman from the University of Turku. This has been a wonderful journey which we have shared together. I am sad to see it end but at the same time I feel like a child on the verge of adulthood: it is time for me to survive on my own! You have given me the best possible basis to build my research career. Elina, I have been so lucky to get to know you and to get you as my supervisor. You have never left me alone and the decisions have been made together. You have challenged my thinking, encouraged me, and supported me with all your heart. Maija, your impact on my PhD process has been remarkable. You never make a big number of yourself but if one listens to your words, one may find out that there is a lot to learn. You have given many insights to my research process and with you, I have always felt that everything is going to be alright. Kirsi, you are the reason why I left on this PhD journey. It all started from a one phone call. You have truly led the way for me in so many ways. There is absolutely no one like you.

I am also deeply grateful to my official reviewers, Docent Hanna-Leena Melender from the University of Oulu, and Professor Tracy Levett-Jones from the University of Technology Sydney, Australia. Thank you for your encouragement and thank you for your excellent and accurate comments which helped me to improve my thesis. Hanna-Leena, I truly appreciate the effort you made to review my thesis. Our discussion was intensive but worth every minute. Tracy, I was so glad when you accepted the invitation to review my thesis. Your comments meant a lot to me. In addition, I would like to thank Professor Helena Leino-Kilpi, the Head of the Department of Nursing Science, for accepting the invitation to be my follow-up committee member. You have made an impact on my scientific thinking since my master's studies in your department. Moreover, I owe my gratitude to Professor Katri

Vehviläinen-Julkunen from the University of Eastern Finland for accepting the invitation to be the official opponent of my dissertation defence.

I wish to thank PhD Janne Engblom and PhLic Eero Laakkonen from the University of Turku for the statistical support and co-writing of the articles. This has been a fruitful co-operation and I have learned a lot. Thank you for your effort, patience, and support on my research process. In addition, I would like to thank Professor James (Jim) Ramsay and Juan Li from the McGill University, Canada, for the consultation concerning the IRT analysis.

I owe my gratitude to the staff, post-doc researchers and doctoral student colleagues at the Department of Nursing Science. I have been lucky to meet many wonderful people there during my studies at the university. I especially owe my gratitude to Professor Leena Salminen. Leena, you have encouraged me to continue my research career, since my master's studies. Thank you for your valuable comments in the seminar group as well. In addition, I wish to thank many wonderful teachers during my doctoral studies, especially Professor Riitta Suhonen, the teacher of the CIDI-course and the Director of the Doctoral Programme in Nursing Science. Furthermore, I am grateful to the members of the Research in Nursing Education (RENE) group. I want to thank Training Expert Anna Mäkinen for the technical support in finalising my thesis. My sincere gratitude goes to my doctoral student colleagues, especially to PhD Anu-Marja Kaihlanen, MNSc Jenni Rinne, MSc Anne Pienimaa, MNSc Kristiina Rosqvist and MNSc Hanna Repo. My very special gratitude goes to MNSc and my dear colleague Mika Alastalo. I am so lucky to have you in my life. Thank you for your peer-support during this journey.

I wish to express my sincere thanks to my colleagues and work community in Laurea University of Applied Sciences. I wish to thank the former director of educational and regional unit, Kati Komulainen and Development Manager, Heli Karjalainen for the support and facilitation given during my PhD studies. Furthermore, I want to thank my wonderful colleagues, especially Riikka Mulder, Emmaculate Tamankag and Irene Latva-Korpela. You have showed interest in my PhD studies, encouraged me, and enabled the progress of my studies by being so flexible in your own way. I also wish to thank my latest work community, Tiina, Marko, Hanna, and Krista in the UAS Student Selection Consortium. Thank you for making my working days such a joy. Otherwise, it would have been impossible to finalise this thesis.

I am deeply grateful to all my near and dear ones. I am lucky to share my life with my lovely family and friends. I am not able to list you all, but please accept my warmest gratitude. Thank you for thinking of me and hoping all the best for my life. My parents, Seija and Jari, thank you for your encouragement and believing in me and my skills. Thank you especially for listening to me and not falling asleep when I have babbled about my PhD studies and research topic which perhaps has not been

your biggest interest. My special thanks go to my big sister, Suvi. Thank you for being my "biggest fan" and loving me so much. Moreover, I wish to thank my sister-in-law, Tiina and her family. Thank you especially for the sleepovers that enabled me to proceed my doctoral courses in Turku. I also wish to thank my dear friends Henrika and Essi. Thank you for helping me to tolerate my stress, thank you for all your support and thank you for celebrating my success with me. My thanks also go to Venla. Thank you for leading the way and "mentoring" my PhD process.

I wish to thank my husband Totte from the bottom of my heart. You are the biggest joy of my life and without you, no doubt, me and my life would be very different. You have given me perspective to look at life more practically, and because you have believed in me, I have believed in myself. Thank you for your love, patience, marriage, and financial support. It is noteworthy, that without you it would have been a lot harder to focus on research and carry on the PhD studies.

I owe a sincere gratitude to all the participants of this doctoral dissertation study. I owe my gratitude to the members of the Reforming Student Selection in Nursing Education (ReSSNE) project.

Finally, I wish to thank everyone who has ever encouraged me in any way during this PhD process. Your kind words and thoughts have lifted my spirit and kept me going.

This study was supported financially by the University of Turku, the Finnish Nursing Education Foundation, and the Finnish Association of Nursing Research.

25 April 2021
Jonna Vierula

References

Alfaro-LeFevre, R. (2013). Critical thinking, clinical reasoning, and clinical judgment: A practical approach (5th ed.). Elsevier Saunders, St. Louis, MO.

American Educational Research Association (AERA); American Psychological Association (APA); National Council on Measurement in Education (NCME) . (2014). Standards for educational and psychological testing. American Educational Research Association. Washington, DC.

Andersson, Å., Frank, C., Willman, A. M., Sandman, P.-O., & Hansebo, G. (2018). Factors contributing to serious adverse events in nursing. Journal of Clinical Nursing, 27(1–2), e354–e362. doi:<https://doi.org/10.1111/jocn.13914>

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19–32. doi:<https://doi.org/10.1080/1364557032000119616>

Armstrong, R., Hall, B. J., Doyle, J., & Waters, E. (2011). "Scoping the scope" of a cochrane review. Journal of Public Health, 33(1), 147–150. doi:<https://doi.org/10.1093/pubmed/fdr015>

Aveyard, H. (2007). Doing a literature review in health & social care: A practical guide. McGraw Hill Companies, Open University Press, Berkshire, UK.

Baker, F. B. (2001). The basics of Item Response Theory (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation, College Park, MD, USA.

Banning, M. (2008). A review of clinical decision making: Models and current research. Journal of Clinical Nursing, 17(2), 187–195. doi:https://doi.org/10.1111/j.1365-2702.2006.01791.x

Brooke, J. (1996). SUS: a 'quick and dirty' usability scale. In P. W. Jordan, B. Thomas, & B. A. Weerdmeester, Usability evaluation in industry (pp. 189–194). Taylor and Francis, London.

Brooke, J. (2013). SUS: A retrospective. .Journal of Usability Studies, 8(2), 29–40. Retrieved January 8, 2021, from <http://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Brooke_February_2013.pdf>

Bucknall, T. K. (2000). Critical care nurses' decision-making activities in the natural clinical setting. Journal of Clinical Nursing, 9(1), 25–36. doi:<https://doi.org/10.1046/j.1365-2702.2000.00333.x>

Burton, R. F. (2005). Multiple-choice and true/false tests: myths and misapprehensions. Assessment & Evaluation in Higher Education, 30(1), 65–72. doi:<https://doi.org/10.1080/0260293042003243904>

Callwood, A., Groothuizen, J. E., Lemanska, A., & Allan, H. (2020). The predictive validity of Multiple Mini Interviews (MMIs) in nursing and midwifery programmes: Year three findings from a cross-discipline cohort study. Nurse Education Today, 88(5), 1–8. doi:<https://doi.org/10.1016/j.nedt.2019.104320>

Cambridge Dictionary. (n.d.). assessment. Retrieved March 12, 2021, from <https://dictionary.cambridge.org/dictionary/english/assessment>

Canova, C., Brogiato, G., Roveron, G., & Zanotti, R. (2016). Changes in decision-making among Italian nurses and nursing students over the last 15 years. Journal of Clinical Nursing, 25(5–6), 811–818. doi:<https://doi.org/10.1111/jocn.13101>

Capponi, N., & Mason Barber, L. A. (2020). Undergraduate nursing program admission criteria: A scoping review of the literature. Nurse Education Today, 92(9), 1–14. doi:<https://doi.org/10.1016/j.nedt.2020.104519>

Carbogim, F. C., Oliveira, L. B., & Püschel, V. A. (2016). Critical thinking: concept analysis from the perspective of Rodger's evolutionary method of concept analysis. Revista Latino-Americana de Enfermagem, 24(e2785). doi:<https://doi.org/10.1590/1518-8345.1191.2785>

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. Contemporary Educational Psychology, 27(2), 270–295. doi:<https://doi.org/10.1006/ceps.2001.1094>

Cerbin, B. (1988). The Nature and Development of Informal Reasoning Skills in College Students. Paper presented at the National Institute on Issues in Teaching and Learning (12th, Chicago, IL, April 24-27, 1988). Retrieved January 7, 2021, from <https://files.eric.ed.gov/fulltext/ED298805.pdf>

Chiu, T.-W., & Camilli, G. (2013). Comment on 3PL IRT Adjustment for Guessing. Applied Psychological Measurement, 37(1), 76–86. doi:<https://doi.org/10.1177/0146621612459369>

Collins, S., & Hewer, I. (2014). The impact of the Bologna process on nursing higher education in Europe: A review. International Journal of Nursing Studies, 51(1), 150–156. doi:<https://doi.org/10.1016/j.ijnurstu.2013.07.005>

Cook, D. A., Erwin, P. J., & Triola, M. M. (2010). Computerized virtual patients in health professions education: A systematic review and meta-analysis. Academic Medicine, 85(10), 1589–1602. doi:<https://doi.org/10.1097/ACM.0b013e3181edfe13>

Crouch, S. J. (2015). Predicting success in nursing programs. Journal of College Teaching & Learning, 12(1), 45–54. doi:<https://doi.org/10.19030/tlc.v12i1.9069>

Dante, A., Valoppi, G., Saiani, L., & Palese, A. (2011). Factors associated with nursing students' academic success or failure: A retrospective Italian multicenter study. Nurse Education Today, 31(1), 59–64. doi:<https://doi.org/10.1016/j.nedt.2010.03.016>

Deschênes, M.-F., Goudreau, J., & Fernandez, N. (2020). Learning strategies used by undergraduate nursing students in the context of a digitial educational strategy based on script concordance: A descriptive study. Nurse Education Today, 95(12), 1–9. doi:<https://doi.org/10.1016/j.nedt.2020.104607>

DeVellis, R. F. (2017). Scale Development: Theory and applications (4th ed.). SAGE Publications. Los Angeles.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., . . . Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. Journal of Nursing Scholarship, 39(2), 155–164. doi:<https://doi.org/10.1111/j.1547-5069.2007.00161.x>

Dimitrov, D., & Shelestak, D. (2003). Psychometric analysis of performance of categories of client needs and nursing process with the NLN diagnostic readiness test. Journal of Nursing Measurement, 11(3), 207–223. doi:<https://doi.org/10.1891/jnum.11.3.207.61270>

Djupsjöbacka, S. (2004). Valinta kokeessa. Valtakunnallinen sosiaali- ja terveysalan opiskelijavalintakokeiden kehittämishanke vuosina 2002–2003. Kartoitusraportti. Diakonia-ammattikorkeakoulun julkaisuja B raportteja 22. Diakonia-ammattikorkeakoulu. Helsinki.

Doody, O., Slevin, E., & Taggart, L. (2013a). Focus group interviews in nursing research: Part 1. British Journal of Nursing, 22(1), 16–19. doi:<https://doi.org/10.12968/bjon.2013.22.1.16>

Doody, O., Slevin, E., & Taggart, L. (2013b). Preparing for and conducting focus groups in nursing research: Part 2. British Journal of Nursing, 22(3), 170–173. doi:<https://doi.org/10.12968/bjon.2013.22.3.170>

Doody, O., Slevin, E., & Taggart, L. (2013c). Focus group interviews. Part 3: analysis. British Journal of Nursing, 22(5), 266–269. doi:<https://doi.org/10.12968/bjon.2013.22.5.266>

Ehrenfeld, M., & Tabak, N. (2000). Value of admission interviews in selecting of undergraduate nursing students. Journal of Nursing Management, 8(2), 101–106. doi:<https://doi.org/10.1046/j.1365-2834.2000.00155.x>

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. Journal of Advanced Nursing, 62(1), 107–115. doi:<https://doi.org/10.1111/j.1365-2648.2007.04569.x>

Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. Educational Leadership, 43(2), 44–48.

Eriksson, E., Korhonen, T., Merasto, M., & Moisio, E.-L. (2015). Sairaanhoitajan ammatillinen osaaminen. Sairaanhoitajakoulutuksen tulevaisuus-hanke. Ammattikorkeakoulujen terveysalan verkosto ja suomen sairaanhoitajaliitto ry. Bookwell Oy. Porvoo.

European Federation of Nurses Associations (EFN). (2015). EFN Guideline for the implementation of Article 31 of the Mutual Recognition of Professional Qualifications Directive 2005/36/EC, amended by Directive 2013/55/EU. EFN Competency Framework. European Federation of Nurses Associations (EFN). Retrieved January 5, 2021, from <http://www.efnweb.be/wp-content/uploads/EFN-Competency-Framework-19-05-2015.pdf>

Eurostat. (2020). Healthcare personnel statistics - nursing and caring professionals. Retrieved January 5, 2021, from <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Healthcare_personnel_statistics_-_nursing_and_caring_professionals&oldid=355980#Healthcare_personnel_.E2.80.94_nursing_professionals>

Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. (2004). An admissions OSCE: The multiple-mini interview. Medical Education, 38(3), 314–326. doi:<https://doi.org/10.1046/j.1365-2923.2004.01776.x>

Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgment: The development of the California Critical Thinking Disposition Inventory. Journal of Nursing Education, 33(8), 345–350. doi:<https://doi.org/10.3928/0148-4834-19941001-05>

Facione, N., Facione, P., & Winterhalter, K. (2011). The Health Sciences Reasoning Test: HSRT test manual vol. 2012. The California Academic Press, Millbrae, CA.

Facione, P. A. (1990). Critical Thinking: A Statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations. American Philosophical Association, Newark, DE.

Finch, M. L., Wilson, D. R., Symonds, K., & Floyd-Tune, K. (2014). Being Interviewed for admission to a BSN program: A Qualitative Inquiry. Advances in Nursing, 2014, 1–5. doi:<https://doi.org/10.1155/2014/310143>

Finnish Institute of Occupational Health (FIOH). (2018). Kognitiivisen ergonomian parantaminen hoitotyössä: Satakunnan sairaanhoitopiirin kehittämishanke. Finnish Institute of Occupational Health. Retrieved January 5, 2021, from <https://www.julkari.fi/bitstream/handle/10024/136179/Kognitiivisen_ergonomian_parantaminen_hoitotyossa.pdf?sequence=1&isAllowed=y>

Finstad, K. (2006). The System Usability Scale and non-native English speakers. Journal of Usability Studies, 1(4), 185–188.

Flinkman, M. (2014). Young registered nurses' intent to leave the profession in Finland Finland – A mixed-method study. University of Turku, Faculty of Medicine, Department of Nursing Science. Annales Universitatis Turkuensis. Academic dissertation.

Francis, R. (2013). Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry: Executive Summary. The Mid Staffordshire NHS Foundation Trust Public Inquiry. The Stationary Office. Retrieved January 5, 2021, from <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/279124/0947.pdf>

Frenk, J., Chen, L., Bhutta, Z. A., Cohen, J., Crisp, N., Evans, T., . . . Zurayk, H. (2010). Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. The Lancet, 376(9756), 1923–1958. doi:<https://doi.org/10.1016/S0140-6736(10)61854-5>

Gale, J., Ooms, A., Grant, R., Paget, K., Marks-Maran, & D. (2016). Student nurse selection and predictability of academic success: The Multiple Mini Interview project. Nurse Education Today, 40(5), 123–127. doi:<https://doi.org/10.1016/j.nedt.2016.01.031>

Gartrell, K., Kent, V., Rock, M., Williams-Cooper, K., Curran, M. S., Durry, A., . . . Mark, H. (2020). Preadmission predictors for first semester course success in a baccalaureate

nursing program. Journal of Professional Nursing, 36(5), 322–329. doi:<https://doi.org/10.1016/j.profnurs.2020.01.007>

Georg, C., K. K., Ulfvarson, J., Jirwe, M., & Welin, E. (2018). A rubric to assess students' clinical reasoning when encountering virtual patients. Journal of Nursing Education, 57(7), 408–415. doi:<https://doi.org/10.3928/01484834-20180618-05>

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. Higher Education, 74(1), 101–114. doi:<https://doi.org/10.1007/s10734-016-0031-y>

Government decree on Universities of Applied Sciences 18.12.2014/1129 (Valtioneuvoston asetus ammattikorkeakouluista). (n.d.). Retrieved January 5, 2021, from <https://www.finlex.fi/fi/laki/ajantasa/2014/20141129>

Grace, S., Orrock, P., Vaughan, B., Blaich, R., & Coutts, R. (2016). Understanding clinical reasoning in osteopathy: a qualitative research approach. Chiropractic & Manual Therapies, 24(6), 1–10. doi:<https://doi.org/10.1186/s12998-016-0087-x>

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Education Today, 24(2), 105–112. doi:<https://doi.org/10.1016/j.nedt.2003.10.001>

Haavisto, E., Hupli, M., Hahtela, N., Heikkilä, A., Huovila, P., E.-L., M., . . . Talman, K. (2019). Structure and content of a new entrance exam to select undergraduate nursing students. International Journal of Nursing Education Scholarship, 16(1), 1–15. doi:<https://doi.org/10.1515/ijnes-2018-0008>

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of Item Response Theory (1st ed.). SAGE Publications, Newbury Park, CA.

Hammond, K. R. (1996). Human judgement and social policy. Irreducible uncertainty, inevitable error, unavoidable justice. Oxford University Press, New York.

Hamshire, C., Jack, K., Forsyth, R., Langan, M., A., & Harris, W. E. (2019). The wicked problem of healthcare student attrition. Nursing Inquiry, 26(3), 1–8. doi:<https://doi.org/10.1111/nin.12294>

Harmon, M. M., & Thompson, C. (2015). Clinical reasoning in pre-licensure nursing students. Teaching and Learning in Nursing, 10(2), 63–70. doi:<https ://doi.org/10.1016/j.teln.2104.12.001>

Harner, A. M. (2014). Components of the Test of Essential Academic Skills as a predictor of first year success in a baccalaureate nursing program. The Faculty of the College of Education. Florida Gulf Coast University. Florida Gulf Coast University. Academic dissertation.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item Response Theory and Health Outcomes Measurement in the 21st Century. Medical Care, 38(Suppl9), II28–II42. doi:<https://doi.org/10.1097/00005650-200009002-00007>

Health Care Professionals Act 559/1994 (Laki terveydenhuollon ammattihenkilöistä). (n.d.). Retrieved January 5, 2021, from <https://finlex.fi/fi/laki/ajantasa/1994/19940559>

Heijltjes, A., van Gog, T., Leppink, J., & Paas, F. (2015). Unraveling the effects of critical thinking instructions, practice, and self-explanation on students' reasoning performance. Instructional Science, 4, 487–506. doi:<https://doi.org/10.1007/s11251-015-9347-8>

Hendricks, S. M., & Krothe, J. S. (2014). Outcomes and lessons earned regarding the use of interviewing for baccalaureate nursing school admission. Journal of Professional Nursing, 30(5), 392–398. Retrieved from <https://doi.org/10.1016/j.profnurs.2014.01.008>

Henriksen, J., Löfmark, A., Wallinvirta, E., Gunnarsdóttir, Þ. J., & Slettebø, Å. (2020). European Union directives and clinical practice in nursing education in the Nordic countries. Nordic Journal of Nursing Research, 40(1), 3–5. doi:<https://doi.org/10.1177/2057158519857045>

Hernandez, M. (2011). Correlating quantitative nursing preadmission variables, ATI test results, and program outcomes including retention, graduation, and licensure. Northern Illinois University, Department of Adult and Higher Education. UMI Dissertation Publishing, ProQuest LLC. Academic dissertation.

Herrera, C. (2012). Student retention in higher education: Examining the patterns of selection, preparation, retention, and graduation of nursing students in the undergraduate pre-licensure nursing program at Arizona State University. Arizona State University. Academic dissertation.

Higgs, J., & Jones, M. (1995). Clinical Reasoning in the Health Professions (2nd ed,). Butterworth Heinemann. Oxford.

Hoffman, K. (2007). Comparison of decision-making by ‘‘expert” and ‘‘novice” nurses in the clinical setting, monitoring patient haemodynamic status post abdominal aortic aneurysm surgery. University of Technology, Sydney. Unpublished Academic dissertation.

Hong, S., & Yu, P. (2017). Comparison of the effectiveness of two styles of case-based learning implemented in lectures for developing nursing students' critical thinking ability: A randomized controlled trial. International Journal of Nursing Studies, 68(3), 16–24. doi:<https://doi.org/10.1016/j.ijnurstu.2016.12.008>

Huang, H.-M., H. C.-Y., Lee-Hsieh, J., & Cheng, S.-F. (2018). Establishing the competences of clinical reasoning for nursing students in Taiwan: From the nurse educators' perspectives. Nurse Education Today, 66(7), 110–116. doi:<https://doi.org/10.1016/j.nedt.2018.04.007>

Huhn, K., Gilliland, S. J., Black, L. L., & Wainwright, S. F. (2019). Clinical reasoning in physical therapy: A concept analysis. Physical Therapy, 99(4), 440–456. doi:<https://doi.org/10.1093/ptj/pzy148>

Imle, M. A., & Atwood, J. R. (1988). Retaining qualitative validity while gaining quantitative reliability and validity: Development of the Transition to Parenthood Concerns Scale. Advances in Nursing Science, 11(1), 61–75. doi:<https://doi.org/10.1097/00012272-198810000-00007>

Itzhaki, M., Hildesheimer, G., Barnoy, S., & Katz, M. (2016). Family involvement in medical decision-making: Perceptions of nursing and psychology students. Nurse Education Today, 40(5), 181–187. doi:<https://doi.org/10.1016/j.nedt.2016.03.002>

Jessee, M. A., & Tanner, C. A. (2016). Pursuing improvement in clinical reasoning: Development of the Clinical Coaching Interactions Inventory. Journal of Nursing Education, 55(9), 495–504. doi:<https://doi.org/10.3928/01484834-20160816-03>

Johansen, M. L., & O'Brien, J. L. (2016). Decision making in nursing practice: A concept analysis. Nursing Forum, 51(1), 40–48. doi:<https://doi.org/10.1111/nuf.12119>

Jokela, T. (2013). P-SUS (positiviinen SUS) -kysely suomeksi: uusi versio (P-SUS (P-SUS [positive SUS] -scale in Finnish: new version). Retrieved January 8, 2021, from <http://hankikaytettavyytta.blogspot.com/2013/05/p-sus-positiviinen-sus-kysely-suomeksi.html>

Kae-Hwa, J., & Gyeong-Ju, A. (2015). Effects of an educational programme on shared decision-making among Korean nurses. International Journal of Nursing Practice, 21(6), 839–846. doi:<https://doi.org/10.1111/ijn.12306>

Kajander-Unkuri, S., Meretoja, R., Katajisto, J., Leino-Kilpi, H., & Suikkala, A. (2020). Students' self-assessed competence levels during nursing education continuum – A cross-sectional survey. International Journal of Nursing Education Scholarship, 17(1), 1–11. doi:<https://doi.org/10.1515/ijnes-2019-0050>

Kajander-Unkuri, S., Salminen, L., Saarikoski, M., Suhonen, R., & Leino-Kilpi, H. (2013). Competence areas of nursing students in Europe. Nurse Education Today, 33(6), 625–632. doi:<https://doi.org/10.1016/j.nedt.2013.01.017>

Kim, W. Y., Shin, Y. J., Lee, J. M., Huh, J. W., Koh, Y., Lim, C.-M., & Hong, S. B. (2015). Modified Early Warning Score Changes Prior to Cardiac Arrest in General Wards. Plos ONE, 10(6), 1–11. doi:<https://doi.org/10.1371/journal.pone.0130523>

Klegeris, A., Barcley McKeown, S., Hurren, H., Spielman, L. J., Stuart, M., & Bahniwal, M. (2017). Dynamics of undergraduate student generic problem-solving skills captured by a campus-wide study. Higher Education, 74(5), 877–896. doi:<https://doi.org/10.1007/s10734-016-0082-0>

Koivisto, J.-M. (2017). Learning clinical reasoning through game-based simulation. Design principles for simulation games. University of Helsinki, Faculty of Educational Sciences. Helsinki Studies in Education. Academic dissertation.

Koivisto, J.-M., Multisilta, J., Niemi, H., Katajisto, J., & Eriksson, E. (2016). Learning by playing: A cross-sectional descriptive study of nursing students' experiences of learning clinical reasoning. Nurse Education Today, 45(10), 22–28. doi:<https://doi.org/10.1016/j.nedt.2016.06.009>

Kowitlawakul, Y., Brenkus, R., & Dugan, N. (2013). Predictors for success for first semester, second-degree Bachelor of Science in Nursing students. International Journal of Nursing Practice, 19(1), 38–43. doi:<https://doi.org/10.1111/ijn.12014>

Lahtinen, P., Leino-Kilpi, H., & Salminen, L. (2014). Nursing education in the European higher education area — Variations in implementation. Nurse Education Today, 34(6), 1040–1047. doi:<https://doi.org/10.1016/j.nedt.2013.09.011>

Lajoie, D. L. (2013). Reading comprehension and nursing education: A missing variable associated with nursing student attrition? University of Wisconsin-Milwaukee. UWM Digital Commons. Academic dissertation.

Lakanmaa, R.-L. (2012). Competence in intensive and critical care nursing – Development of a basic assessment scale for graduating nursing students. University of Turku, Faculty of Medicine, Department of Nursing Science. Annales Universitatis Turkuensis. Academic dissertation.

Lancia, L., Petrucci, C., Giorgi, F., Dante, A., & Cifone, M. G. (2013). Academic success or failure in nursing students: Results of a retrospective observational study. Nurse Education Today, 33(12), 1501–1505. doi:<https://doi.org/10.1016/j.nedt.2013.05.001>

Lauri, S., & Salanterä, S. (2002). Hoitotyön päätöksentekoteoria (Nursing decision-making theory). Hoitotiede, 14(4), 158–166.

Levett-Jones, T., Hoffman, K., Dempsey, J., Yeun-Sim Jeong, S., Noble, D., Norton, C. A., . . . Hickey, N. (2010). The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. Nurse Education Today, 30(6), 515–520. doi:<https://doi.org/10.1016/j.nedt.2009.10.020>

Lewis, E. J., Ludwig, P. M., Nagel, J., & Ames, A. (2019). Student ethical reasoning confidence pre/post an innovative makerspace course: A survey of ethical reasoning. Nurse Education Today, 75(4), 75–79. doi:<https://doi.org/10.1016/j.nedt.2019.01.011>

Li, J., Ramsay, J. O., & Wiberg, M. .. (2019). TestGardener: A Program for Optimal Scoring and Graphical Analysis. In I. M. Wiberg, S. Culpepper, R. Janssen, J. González, & D. (. Molenaar, Quantitative Psychology. IMPS 2017. Springer Proceedings in Mathematics & Statistics, vol 265. Springer, Cham. doi:<https://doi.org/10.1007/978-3-030-01310-3_8>

Liaw, S. Y., Rashasegaran, A., Wong L., F., Deneen, C. C., Cooper, S., Levett-Jones, T., . . . Ignacio, J. (2018). Development and psychometric testing of a Clinical Reasoning Evaluation Simulation Tool (CREST) for assessing nursing students' abilities to recognize and respond to clinical deterioration. Nurse Education Today, 62(3), 74–79. doi:<https://doi.org/10.1016/j.nedt.2017.12.009>

Lievens, F., Patterson, F., Corstjens, J., Martin, S., & Nicholson, S. (2016). Widening access in selection using situational judgement tests: evidence from the UKCAT. Medical Education, 50(6), 624–636. doi:<https://doi.org/10.1111/medu.13060>

Macduff, C., Stephen, A., & Taylor, R. (2016). Decision precision or holistic heuristic?: Insights on on-site selection of student nurses and midwives. Nurse Education in Practice, 16(1), 40–46. doi:<https://doi.org/10.1016/j.nepr.2015.06.008>

Malau-Aduli, B. S., & Zimitat, C. (2012). Peer review improves the quality of MCQ examinations. Assessment & Evaluation in Higher Education, 37(8), 919–931. doi:<https://doi.org/10.1080/02602938.2011.586991>

Mamede, S., de Carvalho-Filho, M. A., de Faria, R. M., Franci, D., Nunes, M. d., . . . Schmidt, H. G. (2020). 'Immunising' physicians against availability bias in diagnostic reasoning: a randomised controlled experiment. BMJ Quality & Safety, 29(7), 550–559. doi:<https://doi.org/10.1136/bmjqs-2019-010079>

McNelis, A. M., Wellman, D. S., Krothe, J. S., Hrisomalos, D. D., McElveen, J. L., & South, R. J. (2010). Revision and evaluation of the Indiana University School of Nursing baccalaureate

admission process. Journal of Professional Nursing, 26(3), 188–195. doi:<https://doi.org/10.1016/j.profnurs.2010.01.003>

Merriam Webster Dictionary. (2021). Reasoning. Retrieved January 7, 2021, from <https://www.merriam-webster.com/dictionary/reasoning>

Ministry of Education and Culture. (2016). Valmiina valintoihin. Ylioppilastutkinnon parempi hyödyntäminen korkeakoulujen opiskelijavalinnoissa. Opetus- ja kulttuuri-ministeriön julkaisuja 2016: 37. Retrieved January 5, 2021, from <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/79291/okm37.pdf>

Ministry of Education and Culture. (2017, August 17). Korkeakoulujen opiskelijavalintojen kehittämisen toimenpiteet 2017-2020. Retrieved January 5, 2021, from <https://minedu.fi/documents/1410845/4154572/Korkeakoulujen+opiskelijavalintojen+kehitt%C3%A4misen+toimenpiteet_20170817.pdf/09af5b53-2658-4866-8a4e-c6aeda33be84>

Ministry of Education and Culture. (2021). Higher education institutions, science agencies, research institutes and other public research organisations. Retrieved January 5, 2021, from <https://minedu.fi/en/heis-and-science-agencies>

Mirza, N. A., Akhtar-Danesh, N., Noesgaard, C., Martin, L., & Staples, E. (2014). A concept analysis of abductive reasoning. Journal of Advanced Nursing, 70(9), 1980–1994. doi:<https://doi.org/10.1111/jan.12379>

Moran-Ellis, J., Alexander, V. D., Cronin, A., Dickinson, M., Fielding, J., Sleney, J., & Thomas, H. (2006). Triangulation and integration: processes, claims and implications. Qualitative Research, 6(1), 45–59. doi:<https://doi.org/10.1177/1468794106058870>

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. Journal of Mixed Methods Research, 1(48), 48–76. doi:<https://doi.org/10.1177/2345678906292462>

Morris, A. (2011). Student standardised testing: Current practices in OECD countries and a literature review. OECD Education Working Papers No. 65. OECD. Retrieved January 6, 2021, from <https://dx.doi.org/10.1787/5kg3rp9qbnr6-en>

Muthén, L. K., & Muthén, B. O. (1998–2012). Mplus user's guide (7th ed.). Muthén & Muthén, Los Angeles.

National Council on Measurement in Education (NCME). (2019). Assessment Glossary. Retrieved January 5, 2021, from <https://www.ncme.org/resources/glossary>

National League for Nursing (NLN). (2015). NLN Nurse Educator Shortage Fact Sheet. Retrieved January 5, 2021, from <http://www.nln.org/docs/default-source/advocacy-public-policy/nurse-faculty-shortage-fact-sheet-pdf.pdf?sfvrsn=0>

Newell, A., & Simon, H. A. (1972). Human problem solving. Prentice-Hall, Englewood Cliffs, NJ.

Newton, S. E., Smith, L. H., Moore, G., & Magnan, M. (2007). Predicting early academic achievement in a baccalaureate nursing program. Journal of Professional Nursing, 23(3), 144–149. doi:<https://doi.org/10.1016/j.profnurs.2006.07.001>

Nursing & Midwifery Council (NMC). (2010). Standards for pre-registration nursing education. Retrieved January 7, 2021, from <https://www.nmc.org.uk/globalassets/sitedocuments/standards/nmc-standards-for-pre-registration-nursing-education.pdf>

Nursing and Midwifery Council (NMC). (2011). Supporting information for implementing NMC standards for pre-registration nursing education. Retrieved January 6, 2021, from <https://www.nmc.org.uk/globalassets/sitedocuments/registration/supporting-information-of-spne-20120629.pdf>

Organisation for Economic Co-operation and Development (OECD). (2019). Health at a Glance 2019: OECD Indicators. OECD Publishing, Paris. Retrieved January 5, 2021, from <https://doi.org/10.1787/4dd50c09-en>

Oxford Cambridge and RSA Examinations (OCR). (2011). Level 2 Award: Thinking and reasoning skills. Oxford Cambridge and RSA Examinations. Retrieved January 7, 2021, from <https://www.ocr.org.uk/Images/80819-specification.pdf>

Perkins, A., Burton, L., Dray, B., Elcock, & K. (2013). Evaluation of a multiple-mini-interview protocol used as a selection tool for entry to an undergraduate nursing programme. Nurse Education Today, 33(5), 465–469. doi:<https://doi.org/10.1016/j.nedt.2012.04.023>

Petersen, K. A., & Lundin, S. (2007). Nursing education in Sweden: recruitment from different socioeconomic backgrounds. Nordic Journal of Nursing Research & Clinical Studies / Vård i Norden, 27(3), 19–23. doi:<https://doi-org.ezproxy.utu.fi/10.1177/010740830702700305>

Petit dit Dariel, O. J., Raby, T., Ravaut, F., & Rothan-Tondeur, M. (2013). Developing the Serious Games potential in nursing education. Nurse Education Today, 33(12), 1569–1575. doi:<https://doi.org/10.1016/j.nedt.2012.12.014>

Pitt, V., Powis, D., Levett-Jones, T., & Hunter, S. (2014). The influence of personal qualities on performance and progression in a pre-registration nursing programme. Nurse Education Today, 34(5), 866-871. doi:<https://doi.org/10.1016/j.nedt.2013.10.011>

Pitt, V., Powis, D., Levett-Jones, T., & Hunter, S. (2015). The influence of critical thinking skills on performance and progression in a pre-registration nursing program. Nurse Education Today, 35(1), 125–131. doi:<https://doi.org/10.1016/j.nedt.2014.08.006>

Plummer-D'Amato, P. (2008). Focus group methodology part 1: Considerations for design. International Journal of Therapy and Rehabilitation, 15(2), 69–73. doi:<https://doi.org/10.12968/ijtr.2008.15.2.28189>

Polit, D. F., & Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. Research in Nursing & Health, 29(5), 489–497. doi:<https://doi.org/10.1002/nur.20147>

Popham, W. J. (1999). Why standardized tests don't measure educational quality. Using Standards and Assessments, 56(6), 8–15.

Ramsay, J. O., Li, J., & Wiberg, M. (2020). Better test scores with TestGardener. An unpublished version. Retrieved January 7, 2021, from <https://www.psych.mcgill.ca/misc/fda/downloads/FDAfuns/OptimalScoreBook.pdf>

Ramsay, J., Wiberg, M., & Li, J. (2019). Full information optimal scoring. Journal of Educational and Behavioral Statistics, 45(3), 0–32. doi:<https://doi.org/10.3102/1076998619885636>

Reinhold, F., Hofer, S., Berkowitz, M., Strohmaier, A., Scheuerer, S., Loch, F., . . . & Reiss, K. (2020). The role of spatial, verbal, numerical, and general reasoning abilities in complex word problem solving for young female and male adults. Mathematics Education Research Journal, 32, 189–211. doi:<https://doi.org/10.1007/s13394-020-00331-0>

Roediger, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5), 1155–1159. doi:<https://doi.org/10.1037/0278-7393.31.5.1155>

Rowley, G. L., & Traub, R. E. (1977). Formula scoring, number-right scoring, and test-taking strategy. Journal of Educational Measurement, 14(1), 15–22. doi:<https://doi.org/10.1111/j.1745-3984.1977.tb00024.x>

Royal College of Nursing. (2012). Quality with compassion: the future of nursing education. Report of the Willis Commission on nursing education. Royal College of Nursing. Retrieved January 5, 2021, from <https://cdn.ps.emap.com/wp-content/uploads/sites/3/2012/11/Willis-Commission-report-2012.pdf>

Sabbag, A., Garfield, J., & Zieffler, A. (2018). Assessing statistical literacy and statistical reasoning: The REALI instrument. StatisticsEducation Research Journal, 17(2), 141–160. Retrieved January 7, 2021, from <http://www.stat.auckland.ac.nz/serj>

Sackett, P. A., Borneman, M. J., & Connelly, B. S. (2008). High-stakes testing in higher education and employment. Appraising the evidence for validity and fairness. American Psychologist, 63(4), 215–227. doi:< https://doi.org/10.1037/0003-066X.63.4.215>

Sadler, J. (2003). Effectiveness of student admission essays in identifying attrition. Nurse Education Today, 23(8), 620–627. doi:<https://doi.org/10.1016/S0260-6917(03)00112-6>

Sairaanhoitajat. (2021). Tilastoja sairaanhoitajista. Retrieved January 5, 2021, from <https://sairaanhoitajat.fi/ammatti-ja-osaaminen/tilastoja-sairaanhoitajista-2/>

SAS Institute Inc. (2015). SAS/SHARE® 9.4: User's Guide (2nd ed.). Retrieved January 8, 2021, from <https://documentation.sas.com/?docsetId=shrref&docsetTarget=titlepage.htm&docsetVersion=9.4&locale=en>

Sauro, J., & Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? Proceedings of the SIGCHI Conference on Human Factors in Computing Systems May 2011, (pp. 2215–2224). doi:<https://doi.org/10.1145/1978942.1979266>

Schmidt, B., & MacWilliams, B. (2011). Admission criteria for undergraduate nursing programs: A systematic review. Nurse Educator, 36(4), 171–174. doi:<https://doi.org/10.1097/NNE.0b013e31821fdb9d>

Schoenfeld, A. H. (2011). How we think: A theory of goal-oriented decision making and its educational applications. Routledge, New York, NY.

Shipman, D., Roa, M., Hooten, J., & Wang, Z. .. (2012). Using the analytic rubric as an evaluation tool in nursing education: The positive and the negative. Nurse Education Today, 32(3), 246–249. doi:<https://doi.org/10.1016/j.nedt.2011.04.007>

Shulruf, B., Bagg, W., Begun, M., Hay, M., Lichtwark, M., Turnock, A., . . . Poole, P. J. (2018). The efficacy of medical student selection tools in Australia and New Zealand. The Medical Journal of Australia, 208(5), 214–218. doi:<https://doi.org/10.5694/mja17.00400>

Shulruf, B., Wang, Y. G., Zhao, Y. J., & Baker, H. (2011). Rethinking the admission criteria to nursing school. Nurse Education Today, 31(8), 727–732. doi:<https://doi.org/10.1016/j.nedt.2010.11.024>

Silén-Lipponen, M., & Korhonen, T. (2020). Osaamisen ja arvioinnin yhtenäistäminen sairaanhoitaja-koulutuksessa – YleSHarviointi-hanke. Kuopio: Savonia-ammattikorkeakoulun julkaisusarja 5/2020, Savonia-ammattikorkeakoulu. doi:<https://www.theseus.fi/bitstream/handle/10024/347289/2020-5yleshArviointi.pdf?sequence=1&isAllowed=y>

Simmons, B. (2010). Clinical reasoning: concept analysis. Journal of Advanced Nursing, 66(5), 1151–1158. doi:<https://doi.org/10.1111/j.1365-2648.2010.05262.x>

Stage, C., & Ögren, G. (2004). The Swdish Scholastic Assessment Test (SweSAT): Development, results and experiences. Retrieved January 8, 2021, from <https://www.umu.se/globalassets/organisation/fakulteter/samfak/institutionen-for-tillampad-utbildningsvetenskap/hogskoleprovet/publications/60585_em-no-49.pdf>

StataCorp. (2017). Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC.

Statistics Finland. (2016). Numbers of polytechnic students and degrees are rising. Retrieved January 5, 2021, from <http://www.stat.fi/til/akop/2015/akop_2015_2016-04-19_tie_001_en.html>

Stenlund, T., Eklöf, H., & Lyrén, P.-E. (2017). Group differences in test-taking behaviour: an example from a high-stakes testing program. Assessment in Education: Principles, Policy & Practice , 24(1), 4-20. doi:<https://doi.org/10.1080/0969594X.2016.1142935>

Stenlund, T., Lyrén, P.-E., & Eklöf, H. (2018). The successful test taker: exploring test-taking behavior profiles through cluster analysis. European Journal of Psychology of Education volume, 33, 403–417. doi:<https://doi.org/10.1007/s10212-017-0332-2>

Stewart, D. W., Shamdasani, P. N., & Rook, D. W. (2007). Focus groups: Theory and practice (2nd ed.). Sage Publications, Thousand Oaks.

Stobart, G., & Eggen, T. (2012). High-stakes testing – value, fairness and consequences. Assessment in Education: Principles, Policy & Practice, 19(1), 1–6. doi:<https://doi.org/10.1080/0969594X.2012.639191>

Stuenkel, D. (2006). At-risk Students: Do theory grades + standardized examinations = success? Nurse Educator, 31(5), 207–212. doi:<https://doi.org/10.1097/00006223-200609000-00007>

Sulis, I., & Toland, M. D. (2017). Introduction to Multilevel Item Response Theory Analysis: Descriptive and Explanatory Models. Journal of Early Adolescence, 37(1), 85–128. doi:<https://doi.org/10.1177/0272431616642328>

Talman, K. (2014). Student selection in nursing education. A follow-up study of two selection methods and their relations to the knowledge, skills and study motivation of nursing students. University of Turku, Faculty of Medicine, Department of Nursing Science. Annales Universitatis Turkuensis. Academic dissertation.

Talman, K., Borodavkin, M., Kanerva, A.-M., & Haavisto, E. (2018b). Ammattikorkeakoulujen uuden digitaalisen valintakokeen kehittäminen – määrittelyvaiheen tulokset. Tutkimusraportti. Metropolia Ammattikorkeakoulun julkaisusarja. Metropolia Ammattikorkeakoulu. Retrieved January 6, 2021, from <http://urn.fi/URN:ISBN:978-952-328-119-6>

Talman, K., Hupli, M., Puukka, P., Leino-Kilpi, H., & Haavisto, E. (2018a). The predictive value of two on-site selection methods of undergraduate nursing students: A cohort study. Journal of Nursing Education and Practice, 8(7), 12–21. doi:<https://doi.org/10.5430/jnep.v8n7p12>

Talman, K., Hupli, M., Rankin, R., Engblom, J., & Haavisto, E. (2020). Emotional intelligence of nursing applicants and factors related to it: A cross-sectional study. 85(2), 104–271. doi:<https://doi.org/10.1016/j.nedt.2019.104271>

Tashakkori, A. M., Johnson, R. B., & Teddlie, C. (2020). Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences (2nd ed.). Thousands Oaks, SAGE Publications, Inc., USA.

Tavakol, M., Rahimi-Madiseh, M., & Dennick, R. (2014). Postexamination analysis of objective tests using the three-parameter Item Response Theory. Journal of Nursing Measurement, 22(1), 94–105. doi:<http://dx.doi.org/10.1891/1061-3749.22.1.94>

Taylor, R., Macduff, C., & Stephen, A. (2014). A national study of selection processes for student nurses and midwives. Nurse Education Today, 34(8), 1155–1160. doi:<https://doi.org/10.1016/j.nedt.2014.04.024>

The European Parliament and the Council of the European Union. (2005). Directive 2005/36/EC of the European Parliament and of the Council. Retrieved January 5, 2021, from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32005L0036&from=EN>

The European Parliament and the Council of the European Union. (2013). Directive 2013/55/EU of the European Parliament and of the Council. Retrieved January 5, 2021, from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0055&from=EN>

The European Parliament and the Council of the European Union. (2016a). Regulation (EU) 2016/679 of the European Parliament and of the Council. Retrieved January 8, 2021, from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN>

The European Parliament and the Council of the European Union. (2016b). Directive (EU) 2016/2102 of The European Parliament and of the Council. Retrieved January 8, 2021, from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L2102&from=EN>

The Finnish Advisory Board on Research Integrity TENK. (2012). Responsible conduct of research and procedures for handling allegations of misconduct in Finland. Finnish Advisory Board on Research Integrity, Helsinki. Retrieved January 5, 2021, from <https://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf>

The Rectors' Conference of Finnish Universities of Applied Sciences Arene. (2020, April 2). Ammattikorkeakouluihin 92 000 hakijaa, kasvua 8 % viime vuodesta. Retrieved January 5, 2021, from <http://www.arene.fi/ajankohtaista/ammattikorkeakouluihin-92-000-hakijaa-kasvua-8-viime-vuodesta/>

Theobald, K. A., & Ramsbotham, J. (2019). Inquiry-based learning and clinical reasoning scaffolds: An action research project to support undergraduate students' learning to 'think like a nurse'. Nurse Education in Practice, 38(7), 59–65. doi:https://doi.org/10.1016/j.nepr.2019.05.018

Thompson, C., Aitken, L., Doran, D., & Dowding, D. (2013). An agenda for clinical decision making and judgement in nursing research and education. International Journal of Nursing Studies, 50(12), 1720–1726. doi:<https://doi.org/10.1016/j.ijnurstu.2013.05.003>

Timer, J. E., & Clauson, M. I. (2011). The use of selective admissions tools to predict students' success in an advanced standing baccalaureate nursing program. Nurse Education Today, 31(6), 601–606. doi:<https://doi.org/10.1016/j.nedt.2010.10.015>

Tower, M., Watson, B., Bourke, A., Tyers, E., & Tin, A. (2019). Situation awareness and the decision-making processes of final-year nursing students. Journal of Clinical Nursing, 28(21–22), 3923–3934. doi:<https://doi.org/10.1111/jocn.14988>

Trobec, I., & Istenic Starcic, A. (2015). Developing nursing ethical competences online versus in the traditional classroom. Nursing Ethics, 22(3), 352–366. doi:<https://doi.org/10.1177/0969733014533241>

Turner, D. W. (2010). Qualitative interview design: A practical guide for novice investigators. The Qualitative Report, 15(3), 754–760. Retrieved January 8, 2021, from <https://nsuworks.nova.edu/tqr/vol15/iss3/19>

Universities and Colleges Admission Service (UCAS). (2020). How to write a UCAS Undergraduate personal statement. Retrieved January 7, 2021, from <https://www.ucas.com/undergraduate/applying-university/how-write-ucas-undergraduate-personal-statement>

Universities of Applied Sciences Act 932/2014 (Ammattikorkeakoululaki). (n.d.). Retrieved January 5, 2021, from <https://finlex.fi/fi/laki/ajantasa/2014/20140932>

University Clinical Aptitude Test (UCAT). (2019). UCAT University Clinical Aptitude Test. Retrieved January 7, 2021, from <https://www.ucat.ac.uk/ucat/practice-tests/>

Valvira National Supervisory Authority for Welfare and Health. (2015). Professional practice rights. Retrieved January 5, 2021, from <https://www.valvira.fi/web/en/healthcare/professional_practice_rights>

Vaughn, S., Schumm, J. S., & Sinagub, J. (1996). Focus group interviews in education and psychology. SAGE, Thousand Oaks, CA.

Vierula, J., Stolt, M., Salminen, L., Leino-Kilpi, H., & Tuomi, J. (2016). Nursing education research in Finland—A review of doctoral dissertations. Nurse Education Today, 37(2), 145–154. doi:<https://doi.org/10.1016/j.nedt.2015.10.014>

Vipunen Education Statistics Finland. (2021a). Students and degrees: Degrees. Retrieved January 5, 2021, from <https://vipunen.fi/en-gb/polytechnic/Pages/Opiskelijat-ja-tutkinnot.aspx>

Vipunen Education Statistics Finland. (2021b). Applicants and selected candidates: Applicants and those who accepted a place in university of applied sciences education. Retrieved January 5, 2021, from https://vipunen.fi/en-gb/polytechnic/Pages/Hakeneet-ja-hyv%C3%A4ksytyt.aspx

Wambuguh, O., Eckfield, M., & Van Hofwegen, L. (2016). Examining the importance of admissions criteria in predicting nursing program success. International Journal of Nursing Education Scolarship, 13(1), 87–96. doi:<https://doi.org/10.1515/ijnes-2015-0088>

Wang, L., Beckett, G. H., & Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. Applied Measurement in Education, 19(4), 305–328. doi:<https://doi.org/10.1207/s15324818ame1904_5>

Waughn, A., Smith, D., Horsburgh, D., & Gray, M. (2014). Towards a values-based person specification for recruitment of compassionate nursing and midwifery candidates: A study of registered and student nurses' and midwives' perceptions of prerequisite attributes and key skills. Nurse Education Today, 34(9), 1190–1195. doi:<https://doi.org/10.1016/j.nedt.2013.12.009>

Webster's Dictionary. (1989). Webster's Unabridged Dictionary of the English Language. Portland House, New York.

Wedman, J. (2017). Theory and validity evidence for a large-scale test for selection to higher education. Umeå University, Department of Applied Educational Science. Educational Measurement. Umeå University. Academic dissertation.

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. Journal of Advanced Nursing, 52(5), 546–553. doi:<https://doi.org/10.1111/j.1365-2648.2005.03621.x>

WHO (World Health Organization). (2009). Nursing & midwifery human resources for health. Global standards for the initial education of professional nurses and midwives. WHO Press, World Health Organization, Geneva. Retrieved January 5, 2021, from <https://www.who.int/hrh/nursing_midwifery/hrh_global_standards_education.pdf>

WHO (World Health Organization). (2013). Transforming and scaling up health professionals' education and training: WHO Education Guidelines 2013. WHO Press, World Health Organization, Geneva. Retrieved January 5, 2021, from <https://apps.who.int/iris/bitstream/handle/10665/93635/9789241506502_eng.pdf;jsessionid=8074B5B6566A340E6D49B14176E74A6F?sequence=1>

WHO (World Health Organization). (2021b). Data and statistics: Shortage of nurses and midwives. Retrieved January 5, 2021, from <https://www.euro.who.int/en/health-topics/Health-systems/nursing-and-midwifery/data-and-statistics>

Wolfe, M. (2017). An assessment of errors and near-misses from pre-licensure student nurses. Retrieved January 7, 2021, from <https://hsrc.himmelfarb.gwu.edu/son_dnp/5>

Wolkowitz, A. A., & Kelley, J. A. (2010). Academic predictors of success in a nursing program. Journal of Nursing Education, 49(9), 498–503. doi:<https://doi.org/10.3928/01484834-20100524-09>

World Health Organization (WHO). (2021a). Health workforce, Nursing & midwifery. Retrieved January 5, 2021, from <https://www.who.int/hrh/nursing_midwifery/en/>

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. Shanghai Archives of Psychiatry, 26(3), 171–177. doi:<https://doi.org/10.3969/j.issn.1002-0829.2014.03.010>

Zuriguel Pérez, E., Lluch Canut, M. T., Falcó Pegueroles, A., Puig Llobet, M., Moreno Arroyo, C., & Roldán Merino, J. (2014). Critical thinking in nursing: Scoping review of the literature. International Journal of Nursing Practice, 21(6), 820–830. doi:<https://doi.org/10.1111/ijn.12347>

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

ISBN 978-951-29-8455-8 (PRINT)
ISBN 978-951-29-8456-5 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)

Painosalama Oy, Turku, Finland 2021