Received: 10 April 2022 | Revised: 7 November 2022 | Accepted: 19 December 2022

DOI: 10.1111/jan.15557

JAN Leading Global Nursing Research WILEY

RESEARCH METHODOLOGY: INSTRUMENT DEVELOPMENT

Development and psychometric evaluation of the Emotional Intelligence Test (EMI-T) for social care and healthcare student selection

Anne Pienimaa[1] | Kirsi Talman[1,2] | Jonna Vierula[3] | Eero Laakkonen[4] | Elina Haavisto[5]

[1]Department of Nursing Science, 20014 University of Turku, Turku, Finland

[2]Metropolia University of Applied Sciences, Helsinki, Finland

[3]UAS Student Selection Consortium, Laurea University of Applied Sciences, Vantaa, Finland

[4]Department of Teacher Education, University of Turku, Turku, Finland

[5]Department of Health Sciences, Tampere University, Tampere University Hospital, Tampere, Finland

Correspondence
Anne Pienimaa, Department of Nursing Science, 20014 University of Turku, Turku, Finland.
Email: anne.m.pienimaa@utu.fi

Funding information
Development Project for Student Selection in Finnish Universities of Applied Sciences (UAS) 2017-2020., Grant/Award Number: OKM/200/523/2016

Abstract

Aim: To develop an emotional intelligence (EI) test and evaluate its psychometrics for social and healthcare student selection.

Design: A cross-sectional methodological design.

Methods: The test was developed based on a systematic review and focus group interviews. Content validity was evaluated with expert panels, and preliminary psychometrics with two pilot studies. Descriptive statistics, correlations and item response theory were used.

Data Sources: Search was conducted in six databases 2018. Focus group interviews were conducted with educators and professionals in 2019. Expert panels with doctoral students, researchers and educators were conducted in 2020. Pilot tests with students were conducted 2020–2021. The developed test was administered to 4808 applicants 2021.

Results: The test included four subscales. Correlations support the test's theoretical structure. The items were mainly easy.

Conclusion: The test assesses EI objectively and comprehensively. The item-level distractor analysis can be used for further test development.

Impacts: Social care and healthcare students engage in clinical practice early in their studies, and these environments can be emotionally challenging. Assessing EI in student selection with adequate test can help the institutions of higher education to select the students with required abilities to succeed in the studies. The assessment of EI during student selection also provides information higher education institutions could use to develop and provide support interventions. The results may also encourage practice placements to include EI elements as learning objective. The results of this study and especially the use of IRT and detailed distractor analysis to evaluate the psychometric properties of EMI-T can benefit researchers and educators that develop or evaluate objective assessment tools with multiple choice questions.

Implications for the profession and/or patient care: Emotional intelligence is important for students to enable professional interaction.

KEYWORDS
emotional intelligence, instrument development, nursing education, psychometric evaluation, social care and healthcare education, student selection

1 | INTRODUCTION

Emotional intelligence (EI) can be defined as a set of abilities including the perception, understanding, expression, management and utilization of emotions (Mayer et al., 2016). EI is important to assess in social care and healthcare student selection (Haavisto et al., 2019; Pienimaa et al., 2022). EI has a positive relationship with academic performance in healthcare students (Pienimaa et al., 2022). In particular, EI seems to improve performance in clinical practice (Lewis et al., 2017; Pienimaa et al., 2022). Social care and healthcare students engage in clinical practice early in their studies, and these environments are emotionally challenging. Abilities such as EI are needed to cope in such environments (Lewis et al., 2017). There is some evidence that healthcare students have higher than average EI (e.g. Aithal et al., 2016). However, a recent study reported significant variations in the EI of nursing applicants (Talman et al., 2020).

Social care and healthcare student selection is a vital topic because it affects numerous institutions of higher education and thousands of applicants worldwide annually. In the United States alone, over 220,000 applicants enrolled in entry-level baccalaureate nursing programmes in the academic year 2018 (American Association of Colleges of Nursing, 2019). Furthermore, the World Health Organization (2021) states that there is currently a lack of skilled health professionals, and it is estimated that there will be a global deficit of more than 7 million nurses and midwives by 2030. It is important for higher education institutions all around the world to select students with sufficient EI so that they can succeed in their studies and graduate on time (Pienimaa et al., 2022). Graduating social care and healthcare students are needed by society to achieve a sufficient workforce.

In Finland, there is a national digital entrance examination (UAS Exam), which purpose is to measure undergraduate social care and healthcare applicants' reasoning, language, mathematical, ethical and EI skills (Vierula, Karihtala, et al., 2021). The purpose of the EI domain is to assess applicants EI and to rank order applicants according to their EI score. The selection decisions are based on overall test performance, but applicant will need to achieve a minimum score from each domain (Vierula, Karihtala, et al., 2021). In Finland, undergraduate (bachelor's degree) social care and healthcare education is provided in Universities of Applied Sciences and the length of degrees varies from 210 to 270 ECTS credits. The expected time of graduation is from 3.5 to 4.5 years depending on the programme (e.g. nursing, social services, physiotherapy, public health nursing, midwifery and paramedic; Ministry of Education and Culture, 2021).

Impact

What already is known?

- Students engage in emotionally challenging placements early in their studies.
- The assessment of applicants' emotional intelligence is suggested.
- Comprehensive and objective selection instruments are missing.

What this article adds?

- Theoretical structure of the developed test is supported.
- Applicants performed well indicating that the test was easy.
- The item response analysis enabled precise psychometric evaluation.

Implications for practice/policy

- The results may help educators to decide what to assess in student selection.
- The description of item-level analysis may benefit educators/researchers in developing objective assessments.
- The results may encourage practice placements to include emotional intelligence as learning objective.

Reporting Method: Strobe

What does this article contribute to the wider global clinical community?

- Assessing emotional intelligence in student selection aims to ensure students' success in studies.

No patient or public contribution:

The test was developed for student selection.

The comprehensive assessment of EI is relevant to ensuring students' abilities to cope with the demands of social care and healthcare studies, and the total EI score seems to be the best predictor of study success (Lewis et al., 2017; Pienimaa et al., 2022). Additionally, higher educational institutions are responsible for fair student selection

processes. The evaluation methods they use must be valid and objective to ensure equitable selection (Haavisto et al., 2019; Talman et al., 2020). However, few of the existing EI instruments can be considered objective and most of them assess only few EI categories. Furthermore, most objective instruments have licensing fees, so financial constraints make it almost impossible to use them for social care and healthcare student selection because of the high number of applicants. The assessment of EI has been previously suggested for the selection phase (Haavisto et al., 2019), but there is a lack of objective and comprehensive instruments and a need for the further operationalization of the concept in this context (Pienimaa et al., 2022; Talman et al., 2020). There are few instruments that can be used to assess EI for social care and healthcare student selection because most instruments have not been validated in these settings (Pienimaa et al., 2022). Furthermore, many objective EI instruments seem to measure only a few categories of EI (e.g. the Situational Test of Emotional Understanding and the Situational Test of Emotional Management; Pienimaa et al., 2022). Due to the lack of comprehensive (i.e. test including several different EI categories), objective and validated EI tests for social care and healthcare settings and especially in the student selection context, there is a need to develop a new objective test for the assessment of EI in the social care and healthcare student selection context.

2 | BACKGROUND

There are several definitions of EI. It can be defined as a set of abilities including the appraisal, expression and regulation of emotions, as well as the ability to use feelings to facilitate performance and solve problems (Mayer et al., 2016). Goleman (1995) introduced a broader definition that includes personality aspects. His definition includes contents such as self-awareness, self-regulation, motivation, empathy and social skills. According to Bar-On (2006), EI can be defined as a collection of personal, emotional and social skills and abilities that affect a person's capacity to successfully cope with environmental pressures and demands. There is some debate about the cultural aspect of EI and whether EI is a stable trait or ability that can be changed in time. Previous research is controversial. Some studies indicate that EI might be affected by cultural background (Johnsen et al., 2012; Zhang & Cross, 2011), although according to the Scherer et al. (2011) there is, evidence for intercultural similarity in perception of emotions. Furthermore, some studies indicate that EI can be improved (Foster et al., 2017; Salminen-Tuomaala, 2020) while others have not noticed improvement in EI during education (Orak et al., 2016) or have found mixed results (i.e. EI has enhanced in some subscales but declined in others) (Shanta & Gargiulo, 2014).

The multiple definitions of EI have led to its different constructions, such as trait EI (Petrides et al., 2007), ability EI (e.g. Mayer et al., 2016) and mixed EI, which refers to both the ability and trait constructs (Bar-On, 2006; Goleman, 1995). Trait EI includes emotion-related behaviours with multi-level personality hierarchies and is typically measured using self-report measures; ability EI includes emotion-related mental skills, such as reasoning validly with emotions and with emotion-related information and is usually measured using objective ability-type measures (Mayer et al., 2016; Petrides et al., 2007).

Most EI instruments are mixed model tools with both ability-based content, such as emotional management, and trait-based content, such as self-awareness (e.g. Emotional Quotient Inventory, Emotional Competence Inventory-University Edition and different versions of Schutte's Emotional Intelligence Test/Scale). These are self-report instruments in which respondents assess their own EI abilities and personalities. Thus, such tools assess respondents' own perceptions and do not give objective evaluations (Bar-On, 2006; Goleman, 1995). Of the existing instruments, the Trait EI Questionnaire Short Form (TeiQue-SF) seems to be the only trait-based tool that has been used to assess the EI of healthcare applicants (Pienimaa et al., 2022). TeiQue-SF assess personality traits, such as self-esteem, self-motivation, empathy, happiness and optimism (Petrides et al., 2007). The most used ability-based EI instrument is the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) (Pienimaa et al., 2022). The MSCEIT requires individuals to solve emotional problems that have correct and incorrect answers rather than to self-report their own perceptions of EI. It assesses EI abilities, such as perception, facilitation, understanding and management of emotions (Mayer et al., 2003).

The process of developing a test requires extensive conseptualization to ensure that the test is assessing comprehensively and distinctly the concept it is supposed to assess. Thus, the concept of EI has been analysed in detail in EI instruments. In a review by Pienimaa et al. (2022), six main EI categories were identified from existing EI instruments that have been used in social care and healthcare selection or education context (Table 1). The EI instruments generally included three to four main EI categories. In a study by Pienimaa et al. (2021), social care and healthcare educators and professionals identified a new EI category to be assessed during student selection—acceptance of emotions. This category is not included in any of the existing EI instruments. In this study, the concept of EI is based on the ability definition of EI (Mayer et al., 2016) and the objective assessment of EI. Trait EI is measured with self-report measurements, and these are not appropriate in student selection where results of the assessment have major impact both to the applicants and higher education institutions (Rankin, 2013). Instead, ability EI is the only construct of EI that can be measured objectively and that is the main reason why ability EI was used in this study. In the student selection context, the evaluation methods must be objective to ensure the fair selection of applicants.

3 | THE STUDY

3.1 | Aim and objective

The aim of this study was to develop an objective EI test (the Emotional Intelligence Test [EMI-T]) and evaluate its psychometric properties for social care and healthcare undergraduate student selection. The ultimate goal was to make social care and healthcare student selection equal and valid and to enable the selection of students

TABLE 1 Conceptualization of EI based on previous EI instruments (modified from Pienimaa et al., 2022)

Main categories of EI	Sub-categories of EI	EI instrument
Perception of emotions	Perception of emotions, perceiving of emotions	AES; SSEIT, MSCEIT
	Emotional self-awareness, attention to feelings, emotion recognition, intrapersonal characteristics; self-awareness, self-confidence	Genos; TMMS; GECo; ECI-U II; EQSAC; EQ-i; TEIQue-SF
	Emotional awareness of others, recognition of emotion in others	Genos; TEIS
Understanding of emotions	Understanding emotion, emotion understanding, emotional understanding	MSCEIT; GECo; STEU
	Clarity of feelings	TMMS
	Emotional appropriateness (ability to differentiate between similar emotions)	TEIS
	Recognize patterns	SEI
Emotional expression	Emotional expression, appraisal and expression of emotions, Intrapersonal characteristics; self-expression	Genos; SSEIT, EQ-i
Emotional management	Regulation of emotions, emotion regulation, emotion management, emotional management/Managing emotions	SSEIT; GECo; MSCEIT; STEM
	Managing own emotions, emotional self-management, emotional self-control, regulation of emotion in the self, self-management, self-control, mood repair, exercise optimism, total mood (self-motivation; happiness, optimism), total mood (self-motivation; happiness, optimism)	AES; SSEIT; Genos; TEIS; TMMS; SEI; EQ-I; TEIQue-SF
	Managing others' emotions, emotional management of others, regulation of Emotion in others	AES; SSEIT; TEIS; Genos; ECI-U II; EQSAC
	Fighting against pressure (emotional management and regulation), stress management	EQ-i; TEIQue-SF
	Coping (change management), adaptability (flexibility, problem solving)	EQ-i; TEIQue-SF
Utilizing emotions	Use of emotions, using emotion, utilizing of emotions, navigate emotions	AES; SSEIT; MSCEIT; SEI
	Emotional reasoning, flexible planning (preference to base life decisions on emotions rather than logic)	Genos; TEIS
	Pursue noble goals	SEI
Social awareness and relations	Interpersonal characteristics (social awareness and interpersonal relationship), social awareness, Relationship management, social competence	TEIQue-SF; ECI-U II; EQSAC
	Empathy (being concerned with and affected by others feelings), increase empathy	TEIS (3); EQSAC; SEI

Abbreviations: AES, Assessing Emotions Scale; ECI-U II, Emotional Competence Inventory-University Edition Version 2; EQ-i, Bar-On Emotional Intelligence Questionnaire; EQSAC, Emotional quotient self-assessment checklist; GECo, The Geneva Emotional Competence Test; Genos, Emotional Intelligence Inventory; MSCEIT, The Mayer-Salovey-Caruso Emotional Intelligence Test; SEI, Six Seconds Emotional Intelligence; SSEIT, The Schutte Self-Report Emotional Intelligence Test; STEM, The Situational Test of Emotional Management; STEU, The Situational Test of Emotional Understanding; TEIS, Tett's Emotional Intelligence scale; TEIQue-SF, Trait EI Questionnaire Short Form; TMMS, Trait Meta-Mood Scale.

with adequate EI skills, which may prevent unnecessary attrition (i.e. dropping out from the degree programme or prolonged studies).

The EMI-T was developed for social care and healthcare student selection because social and healthcare professionals often work together in multi-disciplinary teams to provide quality care to patients and clients. Furthermore, there is no indication that EI differs between social care and healthcare students or professionals (Snowden et al., 2015).

4 | METHODOLOGY

4.1 | Design

This study used a cross-sectional methodological design. The study consisted of two phases: (1) the development of the EMI-T and (2) the psychometric evaluation of the instrument. The scale development process was adapted for both phases (DeVellis, 2017). The development phase included a systematic review, focus group interviews, item generation, expert evaluation and pilot tests. A methodological cross-sectional design was used in the psychometric evaluation of the EMI-T.

The psychometric evaluation was performed using item response theory (IRT) because in the student selection context the most important evidence of the validity is the degree to which the evidence supports the intended interpretation of test scores and their relevance to the proposed use, including specifying the construct the test is intended to measure (American Educational Research Association [AERA], American Psychological Association [APA] and NCME, 2014). Validation of the test's construct can be obtained with empirical evidence such as previous literature, use of the evidence of similar tests and expert judgement (American Educational Research Association [AERA], American Psychological Association [APA] and NCME, 2014) as we have done in this study. Even though the Classical Test Theory (CTT) approach has been applied successfully for many years, it has

some limitations compared with IRT in identifying the item-level evaluation (i.e. item-level difficulty and discrimination) (De Champlain, 2010). From the student selection perspective, the item-level difficulty and function are essential for validity of the test to rank order applicants. Thus, we used IRT approach to achieve a greater understanding of the item-level function and to be able to do full distractor (incorrect response) analysis (Li et al., 2019; Tavakol et al., 2014).

IRT analysis with the TestGardener software (Li et al., 2019) was used in the pilot testing and psychometric evaluation of the EMI-T. IRT is a scarcely applied method in nursing education research (Tavakol et al., 2014). To the best of our knowledge, only one previous study in the field of nursing science has used IRT with TestGardener software (Vierula, Talman, et al., 2021).

The focus of the evaluation in the IRT analysis is on the individual items, enabling the evaluation of different item parameters (DeVellis, 2017; Tavakol et al., 2014) and providing an informative way to analyse composite scales consisting of several categorical items that are summarized into a total score (Wellhagen et al., 2021). The TestGardener software applies modern statistical methods to produce accurate estimates of respondent characteristics using full data and enables item-level analysis and full distractor (incorrect response) analysis (Li et al., 2019; Ramsay et al., 2020). TestGardener can be used to evaluate problems with items and help test developers decide whether to rewrite items to clarify ambiguous wording or to modify incorrect options to make them more plausible. The software provides a visual S-shaped logistic curve (item characteristic curve [ICC]) in which different items' response options can be graphically analysed (Li et al., 2019; Ramsay et al., 2020). It is essential that the items are unambiguous (i.e. clarity of the items and response options) to ensure the applicants' equality. Furthermore, the idea of student selection is to set applicants in rank order, so both the difficulty level of the items is important and the test´s ability to discriminate applicants (i.e. differentiate applicants' skills in the upper ability level; Gierl et al., 2017; Tavakol et al., 2014).

The item analysis in this study was based on a graphical analysis of ICCs with TestGardener (Li et al., 2019), in which an item difficulty, pseudo-guessing and the function of distractors are evaluated visually (Figure 1). Difficulty can be examined by evaluating the shift of the S-shaped curve of the correct response option at different quantiles (i.e. difficult: 75%–95%, moderate to difficult: 50%–75%, easy to moderate: 25%–50%, easy: 5%–25% and very easy: <5%). A shift of the curve to the right indicates a more difficult item (Figure 1; Li et al., 2019; Ramsay et al., 2020).

The pseudo-guessing parameter provides information about the opportunity for low-ability applicants to answer items correctly. Pseudo-guessing can be studied by analysing the starting point of the S-curve along the vertical axis. The higher the starting point, the higher the opportunity of guessing the correct response. The 30% threshold can be considered the cut point for guessing the correct response (i.e. high probability of guessing: >30%; low probability of guessing: ≤30%; Tavakol et al., 2014).

The graphical analysis of ICCs enables a full distractor analysis, in which the curves of all response options can be studied. IRT distractor analysis evaluates whether distractors function properly and are unambiguous (Figure 1; Gierl et al., 2017; Tavakol et al., 2014; Vierula, Talman, et al., 2021).

Figure 1 shows an example ICC of all the response options for one item: the correct response option (blue line), three distractors (red, green and pink lines) and the fifth response option, indicating the applicants who did not respond to the item (orange line). The curve of the correct response option starts from below the 0.3 probability line, indicating that the correct response option is not easy to guess. The curve pierces the *x*-axis between the 5% and 25% quantile, indicating that the item is easy. One of the distractors (red line) functions accordingly, attracting lower ability test-takers. Other distractors are not functional, so most of the responders did not choose them at all.

4.2 | Development of the EMI-T

The EMI-T was developed in 2018–2021 for social care and healthcare student selection purposes. The development phase included two stages: (1) structure and item generation of the EMI-T and (2) content validity evaluation and pilot studies (Figures 2 and 3). The development process included three expert panels and two pilot studies ensuring that, every time the EMI-T was modified, the new or revised items of the test were analysed before further development. The preconditions for the entrance examination were considered in the development process. These preconditions were that the exam was in a digital format, had minimum passing score and multiple choice questions with one correct answer option were used (Figure 2).

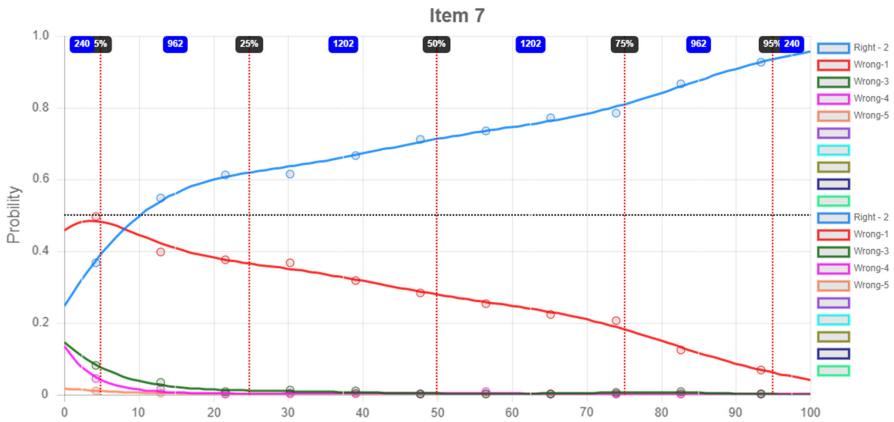

FIGURE 1 Example of a TestGardener item characteristic curve.

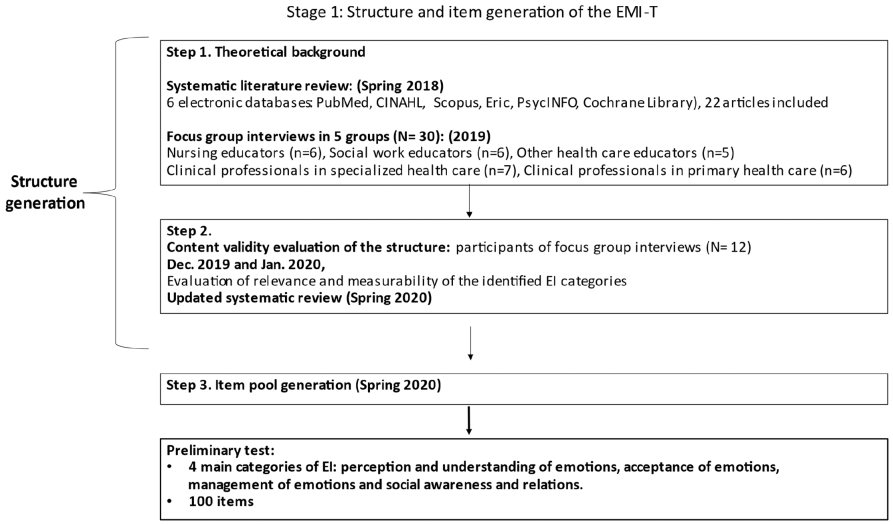

FIGURE 2 Structure and item generation of the EMI-T.

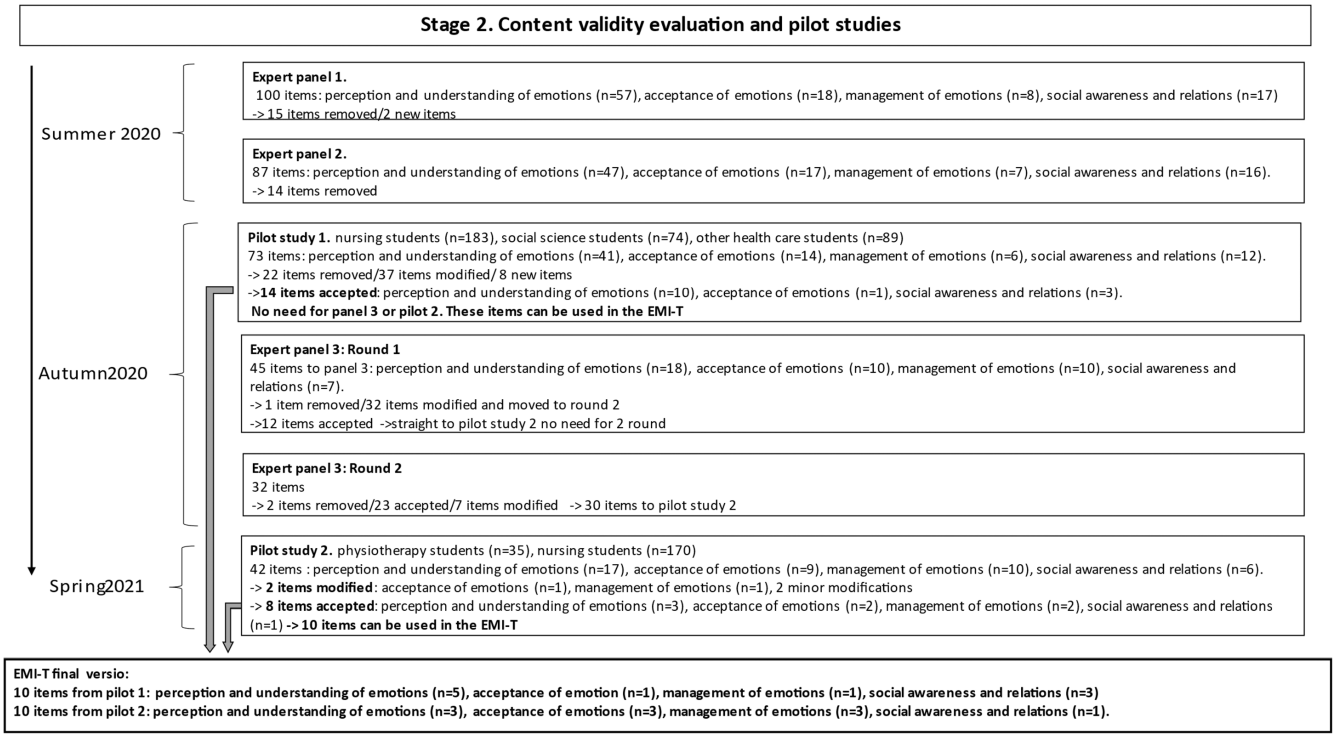

FIGURE 3 Content validity evaluation and pilot studies.

4.2.1 | Stage 1: Structure and item generation of the EMI-T

This stage had three steps: (1) theoretical background (systematic review and focus group interviews), (2) content validity evaluation of the structure (expert panels and updated review) and (3) item pool generation (Figure 1).

Step 1: Theoretical background

The purpose of the theoretical background was to conceptualize EI specifically relevant to social care and healthcare student selection to be able to identify the structure of the EI test. First, a systematic review ($n = 22$) was conducted in 2018 using six electronic databases. Second, five focus group interviews ($n = 30$) with social care and healthcare educators and professionals were conducted in 2019 (Pienimaa et al., 2021; Figure 2). The social care and healthcare educators were recruited via education managers at four universities of applied sciences and clinical professionals were recruited through nurse managers from the country's biggest hospital district and primary healthcare unit. The results of the systematic review were used as themes in the semi-structured focus group interviews. For each theme, the interview questions were as follows: How would you define this theme? What EI content should be assessed when selecting social care and healthcare students? The data were analysed using both deductive and inductive content analysis (Figure 2). The structure of the EMI-T was formed based on the results of the theoretical background.

Step 2: Content validity evaluation of the structure

The content validity of the EI categories identified in step one was evaluated between December 2019 and January 2020. The purpose was to evaluate the relevance and measurability of the identified main EI categories from a student selection perspective. The participants from the focus group interviews (n = 30; social care and healthcare educators and professionals; Pienimaa et al., 2021) were invited for the expert panel. Participants were asked to evaluate the relevance of the EI categories with a 4-point Likert-type scale, and measurability with a dichotomous scale (yes/no). For measurability of the categories, participants were instructed to evaluate if each category would be measurable in a digital entrance exam using objective assessment method with multiple choice questions. Item-related content validity indexes (I-CVIs) were calculated, and the acceptable limit for each category was set to CVI ≥0.78 (Polit & Beck, 2006). After the expert evaluation, the systematic review was updated prior to the generation of the item pool (spring of 2020) to ensure that no relevant new information was missed (e.g. new EI instruments with new categories; Pienimaa et al., 2022; Figure 2).

Step 3. Item pool generation

The item pool was generated in spring 2020 based on the four EI categories. Throughout the study, the items were written by two researchers that were experienced with instrument development and were familiar with the concept of EI. The items were formed following four principles. First, the number of initial items had to be large in anticipation of item reduction during the instrument validation process, although there is no specific standard of how many items should be included to the initial pool (DeVellis, 2017). The aim was to generate many items to each subscale (at least twice the amount that was needed) that deleting several items from each subscale during the development process was not endangering the conceptual adequacy. Second, the items had to be relevant to the content of interest, that is, the items should comprehensively represent the EI categories. Third, the items tasks or questions and response options should be clear and unambiguous (DeVellis, 2017). Fourth, the number of response options had to be optimal to reduce the opportunity of guessing (Gierl et al., 2017).

4.2.2 | Stage 2: Content validity evaluation of the items and pilot studies

This stage had four steps: (1) two consecutive expert panels, (2) first pilot study, (3) third expert panel with two rounds and (4) second pilot study. The content validity of the items was evaluated with expert panels to ensure item clarity, relevancy and representativeness. Two pilot studies were used to undertake a preliminary item analysis and gain an understanding of the functionality and difficulty levels of the developed items (Figure 3).

First, two expert panels were conducted in June 2020 and August 2020. The first panel involved doctoral students (n = 6) and postdoctoral researchers (n = 5) who had knowledge of instrument development (Figure 3). The second panel involved experienced social care and healthcare educators (n = 6) who had an average of 18 years of working experience (5–34 years); almost all (five out of six) had experience with entrance examination organization, development or evaluation. Both expert panels evaluated the clarity and validity of the items. They also evaluated whether they agreed with the correct options. The experts were asked to make suggestions on how the items might be improved. The I-CVIs were calculated and items with I-CVI under 0.78 were rejected (Polit & Beck, 2006; Figure 3).

The first pilot study was conducted in September 2020. The participants were first year social care and healthcare students from two universities of applied sciences who started their studies in August 2020 (n = 346) (Figure 3). They were mainly female (82.7%) and young adults (mean age: 22 years, age range: 18–54); over half were nursing students (52.9%) and had a previous degree (59.4%), often from practical nursing (36.4%). They answered the online version of the 73-item EMI-T version during class. Descriptive statistics were used to overview the data. Frequencies and percentages were calculated for each item. IRT analysis was conducted with TestGardener, and a graphical analysis of ICCs was conducted. The difficulty of items and the functionality of distractors were examined (Figure 3).

The third expert panel was undertaken at the end of 2020. It included two rounds and evaluated those items (n = 45) that were modified based on the results of the first pilot study. This expert panel included experienced social care and healthcare educators (n = 10: round 1; n = 8: round 2) who also had experience in student selection. The experts evaluated the clarity of the items and whether they agreed with the correct options. In the first round, they evaluated the difficulty of the items (4-point Likert scale: 1 = easy, 4 = difficult). Because of a technical error, information for the I-CVIs was not obtained from round one. Thus, in this round, items were evaluated according to the experts' difficulty evaluation and comments. In second round, the I-CVIs on clarity and agreement for the correct options were calculated, and comments were considered. Items with I-CVIs <0.75 were rejected. In round two, the experts did not evaluate the difficulty level (Figure 3).

The second pilot study was conducted in January 2021 to evaluate the modified items (n = 42) from the third expert panel. This study included healthcare students from two universities of applied sciences who started their studies in January 2020 (n = 205). The participants were mainly female (81.1%) and young adults (mean age: 26.4 years, range: 19–54), and over half had a previous degree (68%), which was often practical nursing (35.7%). Descriptive statistics was used to overview the data. Frequencies and percentages were calculated for each item. The frequencies of all item options were overviewed in simple items to obtain more information about the selection of distractors and strengthen the decision to exclude poor distractors. IRT analysis was conducted with TestGardener, and a graphical analysis of ICCs was conducted. The difficulty of the items, pseudo-guessing level and the functionality of the distractors were evaluated (Figure 3).

4.3 | Psychometric evaluation of the EMI-T

After the development process, the EMI-T was psychometrically evaluated with IRT method. Data were collected using the EMI-T from undergraduate social care and healthcare applicants from 20 universities of applied sciences who gave their consent to participate in this study and took the digital entrance examination on 2 days, 31 May and 4 June 2021 ($N = 4808$). The applicants received information about the study during the application process. Informed consent was obtained from the participants before the start of the digital entrance examination. Most of the participants were female and had a high school or vocational school education (Table 2).

First, descriptive statistics (frequencies, percentages, means and standard deviations [SDs]) were calculated for the participant demographics, each item (i.e. frequency and percentage of applicants who got the item correctly or incorrectly) and the total EMI-T score. The item analysis of the EMI-T was based on Pearson's correlation coefficients for item level (correlations between EI subscales, correlations between EI subscales and the total score, and individual items' correlations with one another) and IRT, including the graphical analysis of the ICCs with TestGardener. Graphical analysis included the evaluation of difficulty, pseudo-guessing parameters and the functionality of distractors. The data were analysed using Statistical Analysis Software (SAS 9.4®; SAS Institute Inc., 2015) and TestGardener (Li et al., 2019; software online version).

4.4 | Ethical considerations

This study followed the ethical principles specified by the National Advisory Board on Research Ethics (Finnish Advisory Board on Research Integrity, 2012). Permission to undertake the study was obtained from the institutions involved, and ethics approvals were sought from the ethics committee of the higher education institutions (for expert panels and pilot studies: 10 June 2020; for the psychometric evaluation of the EMI-T during the entrance examination: 14 May 2021). The participants were informed about their anonymity, their right to withdraw from the research and the voluntary nature of the study. Informed consent was obtained from the participants at all stages of the study.

TABLE 2 Demographic information of the applicants ($N = 4808$)

Demographic information	*N*	%
Age (years)		
Under 20	712	14.8
20–24	1772	36.9
25–29	824	17.1
Over 29	1500	31.2
Gender		
Female	3862	80.3
Male	946	19.7
Previous education		
High school	2134	44.4
Vocational school	1920	39.9
Higher education	440	9.2
Other	314	6.5

5 | RESULTS

5.1 | Development of the EMI-T

5.1.1 | Stage 1: Structure and item generation of the EMI-T

Step 1: Theoretical background

In the systematic review, 18 existing EI instruments that have been used in social care and healthcare selection or education context were analysed to conceptualize the EI. As a result, six EI categories (perception of emotions, understanding of emotions, emotional expression, managing emotions, utilizing emotions and social awareness and relations) were identified (Pienimaa et al., 2022). Furthermore, five focus group interviews ($n = 30$) with social care and healthcare educators and professionals were conducted in 2019 (Pienimaa et al., 2021). The content analysis of the focus group interviews yielded a new EI category: acceptance of emotions. Furthermore, also the description of other EI categories was elaborated especially from the student selection perspective. Thus, the systematic review and focus group interviews yielded seven main categories of EI: perception of emotions, understanding of emotions, acceptance of emotions, management of emotions, emotional expression, utilization of emotions and social awareness and relations (Pienimaa et al., 2021).

Step 2: Content validity evaluation of the structure

The content validity of the EI categories identified in step one was evaluated between December 2019 and January 2020. The participants from the focus group interviews (step 1) were invited for the expert panel. The response rate was 40% (12/30).

According to the results, perception of emotions, understanding of emotions, acceptance of emotions, management of emotions and social awareness and relations were the most relevant and measurable categories of EI during social care and healthcare student selection. The EI categories of emotional expression and utilization of emotions were excluded from the test at this point because the I-CVI was under acceptable level 0.78 in these categories. Furthermore, the experts stated that objective evaluation of emotional expression during the digital entrance examination would not be possible, and that utilization of emotions is an ability that will most probably enhance during the education and this ability is not something that necessarily should be assessed in the student selection process. The categories of perception and understanding of emotions were combined based on the comments of the panel members about the difficulties in measuring these two categories separately. It was unclear

when the perception of emotions develops into the understanding of emotions. All in all, four EI categories were identified as relevant from the student selection perspective. The systematic review was updated prior to the generation of the item pool (spring of 2020) to ensure that no relevant new information was missed. Although six new studies were identified, the results did not yield to any new categories.

Step 3. Item pool generation

After the structure and item generation of the EMI-T, its first version included 100 items in four subscales: 57 items in perception and understanding of emotions, 18 in acceptance of emotions, eight in management of emotions and 17 in social awareness and relations (Figure 3). All the items in the preliminary version included one correct response option and three to five distractors.

Stage 2: Content validity evaluation of the items and pilot studies

First, two expert panels were conducted in June 2020 and August 2020. According to the results of these expert panels in 64 items, some minor changes were made to improve clarity. Altogether, 27 items were removed, and two new items were generated (Figure 3).

The first pilot study was conducted in September 2020. Based on the descriptive statistics, over half of the items ($n = 40$) were extremely easy (>95% of the respondents chose the correct options). With these extremely easy items, the frequencies of all item options were overviewed to obtain more information about the selection of distractors and to strengthen the decision to exclude poor distractors. According to the graphical analysis of the ICCs, those items that seemed to be extremely easy—those whose distractors were not selected at all and for which the probability to guess the correct option was extremely high—were excluded (21 items). However, 19 easy items—in which applicants had selected distractors and for which the probability for guessing was not extremely high (over 0.8)—were further modified to increase the difficulty level. They were further evaluated in the third expert panel and the second pilot study. Some of the items ($n = 34$) had six response options, and these were reduced to four (i.e. those poor distractors that only a few respondents chose were deleted) to make the items more difficult. This item reduction brought the response options closer to each other and standardized the response options in all items, that is, all items included four response options. Furthermore, in 37 items, the assignment and/or response options were modified or changed to enhance clarity and difficulty and to reduce guessing (Figure 3).

The third expert panel was undertaken at the end of 2020. It included two rounds and evaluated those items ($n = 45$) that were modified based on the results of the first pilot study. In first round, items that were estimated to be easy and had no comments were considered suitable for the pilot study ($n = 12$). All other items were sent to round two ($n = 32$). After this round, one correct option was modified, and three items were removed (Figure 3).

The second pilot study was conducted in January 2021 to evaluate the modified items ($n = 42$) from the third expert panel. Some of the items ($n = 8$) were extremely easy (>95% of the respondents chose the correct options), and in 12 items, between 90.78% and 94.18% of the respondents chose the correct options. The IRT method was not applicable to these items because of the small sample size and the high percentage of correct responses. These 20 items were removed prior to the IRT analysis, which included 22 items. Those items that seemed to be extremely easy and for which the probability of guessing the correct answer was high were excluded ($n = 8$). Furthermore, items that were not unambiguous were removed (i.e. distractor was chosen as correct answer more often than the correct answer even with the high ability responders) ($n = 6$). Overall, 10 items from the first pilot study, and 10 items from the second pilot study were included in the final EMI-T (Figure 3).

5.2 | Psychometric evaluation of the EMI-T

5.2.1 | The Emotional Intelligence Test)

The EMI-T included 20 multiple choice items (case-based questions or questions related to facial expressions) in four subscales: perception and understanding of emotions (eight items), acceptance of emotions (four items), management of emotions (four items) and social awareness and relations (four items). The subscale of perception and understanding of emotions combined two EI categories, and thus, this subscale had twice as many items than the other subscales (Figure 3). Each item included one correct response option and three incorrect ones. Each correct response yielded one point. Negative points were not used. The minimum score was set to five points based on the fact that an applicant with zero EI ability cannot be selected. The EMI-T included eight pictures with facial expression and applicants needed to recognize different universal emotions from the pictures. Here is also a fictional example of the multiple choice item: Sara has met her best friend. How does she most probably feel? (1) Frustrated, (2) Happy, (3) Sad and (4) Angry.

5.2.2 | Descriptive results and correlations

The mean total score for the EMI-T was 15.92 (SD: 2.16), and the median was 16. The total score range was 0–20. Two-thirds of the applicants achieved a score of 17.00. Less than 1% ($n = 19$) of the applicants failed to receive the minimum passing score (five points). Thus, most of the applicants performed well in the EMI-T.

There was a positive and statistically significant correlation between subscales and between subscales and the total score. The correlation estimates were classified from negligible to very strong (Schober et al., 2018). All subscales had a moderate to strong correlation with the total EMI-T score ($r = 0.60–0.74$) but had a lower correlation with one another ($r = 0.18–0.32$) (Table 3). The correlations between items were weak ($r = -0.21 - 0.27$). The item-to-total score correlations were positive ($r = 0.18–0.40$) and statistically significant ($p < 0.05$; Schober et al., 2018).

TABLE 3 Pearson correlation coefficients (*r*) for the subscales and total scores

	1	2	3	4	Total score
1	1.00[d]	0.22[a]	0.18[a]	0.20[a]	0.74[c]
2	0.22[a]	1.00[d]	0.30[a]	0.30[a]	0.65[b]
3	0.18[a]	0.30[a]	1.00[d]	0.32[a]	0.60[b]
4	0.20[a]	0.30[a]	0.32[a]	1.00[d]	0.62[b]
Total score	0.74[c]	0.65[b]	0.60[b]	0.62[b]	1.00[d]

Note: Subscales: 1 = perception and understanding of emotions, 2 = acceptance of emotions, 3 = management of emotions and 4 = social awareness and relations.

$p < 0.05$.

[a] 0.10–0.39 Weak correlation.

[b] 0.40–0.69 Moderate correlation.

[c] 0.70–0.89 Strong correlation.

[d] 0.90–1.00 Very strong correlation.

The correlations between subscales supported the theoretical structure of the test, indicating that EI is based on different categories that have a significant correlation with one another but an even stronger correlation with the total score.

5.3 | IRT analysis

The evaluation of the difficulty of the items showed that most of them proved to be very easy ($n = 13$) or easy ($n = 4$) (Table 4), indicating that either the correct response options were too obvious or that the incorrect response options failed to be functional distractors. Furthermore, over half of the items (13/20) were susceptible to guessing. Altogether, only five items were fully functional. They included functional distractors and did not exceed the 30% threshold for guessing the correct response (Tables 4 and S4).

TABLE 4 IRT item difficulty and pseudo-guessing levels visually evaluated from the ICC plots

	Correct answer % (*n*)	Item difficulty levels[a]	Pseudo-guessing levels[b]
Perception and understanding of emotions			
Item 1	75.6 (3634)	0	1
Item 2	91.1 (4352)	0	1
Item 3	36.2 (1726)	4	0
Item 4	72.4 (3473)	1	1
Item 5	44.9 (2139)	3	0
Item 6	85.8 (4121)	0	1
Item 7	70.5 (3383)	1	0
Item 8	59.2 (2833)	2	1
Acceptance of emotions			
Item 9	94.5 (4538)	0	1
Item 10	86.3 (4134)	0	1
Item 11	93.5 (4473)	0	1
Item 12	67.9 (3247)	1	0
Management of emotions			
Item 13	94.2 (4518)	0	1
Item 14	85.9 (4112)	0	1
Item 15	89.5 (4281)	0	1
Item 16[c]	95.3 (4582)	N/A	N/A
Social awareness and relations			
Item 17[c]	99.1 (4765)	N/A	N/A
Item 18	76.4 (3638)	1	0
Item 19	92.00 (4397)	0	1
Item 20	88.6 (4226)	0	1

[a]Difficulty levels: 4 = Difficult (75%–95%), 3 = Moderate to difficult (50%–75%), 2 = Easy to moderate (25%–50%), 1 = Easy (5%–25%) and 0 = Very easy (<5%).

[b]Pseudo-guessing: 1 = High (>30%), 2 = Low (≤30%).

[c]Items 16 and 17 were extremely easy (less than 5% chose the distractors), the items were not estimated properly and IRT was not applicable.

6 | DISCUSSION

This study aimed to develop and evaluate the psychometric properties of the EMI-T to assess EI in the social care and healthcare undergraduate student selection context. EI is important to determine in healthcare student selection (Haavisto et al., 2019; Pienimaa et al., 2022), but there has been a lack of tests to assess it objectively and comprehensively in this context (Pienimaa et al., 2022). A fair and comprehensive assessment is crucial to ensure an equitable selection process (Talman et al., 2020). The development process of the EMI-T was versatile, including several phases (two stages and 10 steps), methods and content validity evaluation (three expert panels and two pilot studies). The EMI-T was based on the most relevant and objectively measurable EI categories to be assessed in social care and healthcare student selection. It is comprehensive, including four EI categories that have been incorporated into previous EI instruments (Mayer et al., 2003; Pienimaa et al., 2022), and it also has a new category, acceptance of emotions, which is not included in any of the previous EI instruments (Pienimaa et al., 2021).

The psychometrical testing was based on descriptive statistics, correlations and graphical analysis of the ICCs with TestGardener, in which item-level analysis is possible, enabling more detailed information of the items and full distractor analysis (Gierl et al., 2017; Li et al., 2019; Tavakol et al., 2014). In the psychometric evaluation, the sample size was large enough for statistical analysis and larger than those of other similar student selection studies or studies evaluating EI instruments' psychometric properties. The sample represented the typical characteristics of the population. Previously, IRT with graphical analysis of the ICCs using TestGardener is a scarcely applied method in nursing sciences. The graphical interpretation proved to be excellent in facilitating item-level analysis, including precise distractor analysis. The

preliminary psychometric results during the development stage provided support for the content validity of the developed test, although in the psychometric testing the EMI-T proved to be easy. There was also a positive and statistically significant correlation between subscales and between subscales and the total score, which supports the theoretical structure of the EMI-T. The mean total score for the EMI-T (15.92/20 points) and the IRT analysis indicated that most of the items were easy and that most of the incorrect response options failed at being functional distractors, so the test failed to be very discriminative (i.e. ability to differentiate applicants' skills in the upper ability level). However, the mean total score was under 16 points and SD was 2.16, indicating that the EMI-T still discriminates at the upper ability level and that all applicants did not get maximum scores. Although very easy items are usually removed from the test, it might still be valid to include easy ones to maintain content coverage and ensure that the content is comprehensively measured with the test (Gierl et al., 2017). Including some easy items in the EMI-T was necessary to provide a comprehensive assessment of EI, including all relevant EI categories. Assessing EI extensively is relevant to ensuring students' abilities to cope with the demands of social care and healthcare studies (Lewis et al., 2017; Pienimaa et al., 2022).

Previous research shows that healthcare students seem to have higher than average EI (e.g. Aithal et al., 2016), and the minimum score for the test is currently set to be relatively low (5 points out of 20), so these factors might at least partly explain why the applicants scored high in the EMI-T. Furthermore, the aim of the EMI-T is to assess whether applicants possess reasonable EI to cope with emotional situations during their studies and be academically successful (Lewis et al., 2017; Pienimaa et al., 2022). The purpose of the EMI-T is to select not only those applicants who have tremendous EI but also those with adequate EI to be able to cope with the emotional demands of social care and healthcare studies.

The minimum score of the test is relatively low. This might require further consideration, and more research is needed to set the minimum score at the right level. There is no consensus on the optimal EI level (Davis & Nichols, 2016), although Li et al. (2015) implicated that moderate EI in nursing students could be associated with best coping in possible adverse situations. It is also argued that higher EI is not necessarily always better (Davis & Nichols, 2016). According to a review by Davis and Nichols (2016), those with high EI ability may have enhanced reactivity to stress, and high levels of EI might be related to poorer psychological health. Applicants EI scores at admission could be used to survey which kind of support the students admitted to the programme will need to cope with the emotional demands of the studies.

As stated earlier, incorrect responses mostly failed as functional distractors, which reflected the poor quality of these alternatives and yielded poor discrimination. The revision of the dysfunctional distractors and their further testing are needed to increase the difficulty level of the items. For example, in those items in which pseudo-guessing has been probably and most of the applicants have chosen the correct response (i.e. the items are also easy), the distractors should be more difficult so that these could function better as distractors. In such items, the improvement should concentrate on making the distractors more difficult and less obvious. However, in those items in which pseudo-guessing has been probably but the items are not easy (i.e. distractors have been functioning properly), it is not enough to concentrate on making the distractors more difficult. The layout and wording of the assignment in these items should also be re-examined.

6.1 | Strength and limitations of the work

The strength of this study was several different and versatile methods used in the development process to ensure the content validity and psychometric validity of the EMI-T (i.e. systematic, review, several focus group interviews and expert panels and two pilot studies). Furthermore, in psychometric testing, the sample size was big enough for statistical analysis and large in comparison with sample sizes in other similar student selection studies or studies testing EI instruments' psychometric properties. This study had some limitations concerning the expert panels, the participants of the pilot studies and the psychometric evaluation. This study included three expert panels during the development process, including two rounds in panel three, following the recommended use of this method (Polit & Beck, 2006). However, the first two expert panels received a test version that included 73–100 items. The evaluation form was long and time consuming to fill out. This might have affected the quality of the items if the experts experienced the evaluation strenuous and did not have strength to comment the items at the end part of the form.

In the pilot studies, the study population included students who had just started their education prior to the pilot study. Thus, they represent applicants who have been selected for the programmes. Furthermore, the participants in the pilot studies took the EMI-T in class, but there was no such high-stakes situation in the entrance examination; this might have affected their responses compared with their answers during the actual entrance examination. IRT analysis was successfully performed, but several of the items during the development process were so easy (>95% of the participants chose the correct answers) that the IRT method was not applicable. Thus, a full psychometric evaluation was not possible for these items. Furthermore, the fact that EMI-T included several easy items was a limitation.

6.2 | Recommendations for further research

Further development and psychometric evaluation is needed to enhance EMI-T's difficulty and discrimination levels. Previously, IRT with graphical analysis of ICCs using TestGardener is a scarcely applied method in the nursing sciences. The graphical analysis proved to be excellent in facilitating item-level analysis, including precise distractor analysis. The results can be used for

further revision of the test, especially in relation to improving the distractor items, and thus, the difficulty and discrimination levels of the instrument. The predictive validity of the test should be evaluated in future research.

7 | CONCLUSION

The results of this study and especially the use of IRT and detailed distractor analysis during both the development process and an evaluation of the psychometric properties of EMI-T can benefit researchers and educators that develop or evaluate objective assessment tools with multiple choice questions. In the future, a follow-up study investigating whether applicants EI scores at admission have an impact on study success should be done. The results of this study can also be used to further development of the EMI-T for student selection purposes. Especially, the optimal cut-off level should be further defined. Social care and healthcare students engage in clinical practice early in their studies, and these environments are emotionally challenging. Assessing EI in student selection provides information that higher education institutions could use to develop and provide support interventions for those students that might have difficulties in EI. Furthermore, the results may encourage practice placements to include EI elements as learning objectives.

AUTHOR CONTRIBUTIONS

Anne Pienimaa: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, validation, visualization, writing—original draft and writing—review and editing. Kirsi Talman: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, validation, visualization, writing—original draft and writing—review and editing. Jonna Vierula: Conceptualization, formal analysis, validation, visualization and writing—review and editing. Eero Laakkonen: Data curation, formal analysis, funding acquisition, validation, visualization and writing –review and editing. Elina Haavisto: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, validation, visualization, Writing—original draft and writing—review and editing. All authors participated in commenting on and revising the article critically for important intellectual content and agreed on the final version.

ACKNOWLEDGEMENTS

The authors thank all the participants and institutions that participated to this study and gave their contribution. This research received funding for statistical analysis from Development Project for Student Selection in Finnish Universities of Applied Sciences (UAS) 2017–2020 (Grant Number: OKM/200/523/2016). There is a statistician on the author team (Eero Laakkonen).

CONFLICT OF INTEREST

No conflict of interest has been declared by the authors.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jan.15557.

DATA AVAILABILITY STATEMENT

Author elects to not share data.

IMPACT STATEMENT

Social care and healthcare students engage in clinical practice early in their studies, and these environments can be emotionally challenging. Assessing EI in student selection with adequate test can help the institutions of higher education to select the students with required abilities to succeed in the studies. The assessment of EI during student selection also provides information higher education institutions could use to develop and provide support interventions. The results may also encourage practice placements to include EI elements as learning objective. The results of this study and especially the use of IRT and detailed distractor analysis to evaluate the psychometric properties of EMI-T can benefit researchers and educators that develop or evaluate objective assessment tools with multiple choice questions.

ORCID

Anne Pienimaa https://orcid.org/0000-0003-0857-5776
Jonna Vierula https://orcid.org/0000-0002-7538-9837

TWITTER

Anne Pienimaa @APienimaa

REFERENCES

Aithal, A. P., Kumar, N., Gunasegeran, P., Sundaram, S. M., Rong, L. Z., & Prabhu, S. P. (2016). A survey-based study of emotional intelligence as it relates to gender and academic performance of medical students. *Education for Health*, *29*, 255–258. https://doi.org/10.4103/1357-6283.204227

American Association of Colleges of Nursing. (2019). Annual Report 2019. https://www.aacnnursing.org/Portals/42/Publications/Annual-Reports/Annual-Report-2019.pdf

American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME]. (2014). *Standards for educational and psychological testing* (2014th ed.). American Educational Research Association.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, *18*, 13–25.

Davis, S. K., & Nichols, R. (2016). Does emotional intelligence have a "dark" side? A review of the literature. *Frontiers in Psychology*, *7*, 1316. https://doi.org/10.3389/fpsyg.2016.01316

De Champlain, A. F. (2010). A primer on classical test theory and item response theory for assessments in medical education. *Medical Education*, *44*, 109–117. https://doi.org/10.1016/j.nedt.2017.05.00810.1111/j.1365-2923.2009.03425.x

DeVellis, R. F. (2017). Scale development: Theory and applications. In *Applied social research methods* (4th ed.). SAGE.

Finnish Advisory Board on Research Integrity (TENK). (2012). *Responsible conduct of research and procedures for handling allegations of misconduct in Finland*. http://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf.%20

Foster, K., Fethney, J., McKenzie, H., Fisher, M., Harkness, E., & Kozlowski, D. (2017). Emotional intelligence increases over time: A longitudinal study of Australian pre-registration nursing students. *Nurse Education Today*, *55*, 65–70. https://doi.org/10.1016/j.nedt.2017.05.008

Gierl, M. J., Bulut, O., Qi Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research*, *87*, 1082–1116. https://doi.org/10.3102/0034654317726529

Goleman, D. (1995). *Emotional intelligence. Why it can matter more than IQ*. Bantam Books.

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

Johnsen, B. H., Meeùs, P., Meling, J., Rogde, T., Eid, J., Esepevik, R., Olsen, O. K., & Sommerfelt-Pettersen, J. (2012). Cultural differences in emotional intelligence among top officers on board merchant ships. *International Maritime Health*, *63*(2), 90–95.

Lewis, G. M., Neville, C., & Ashkanasy, N. M. (2017). Emotional intelligence and affective events in nurse education: A narrative review. *Nurse Education Today*, *53*, 34–40. https://doi.org/10.1016/j.nedt.2017.04.001

Li, J., Ramsay, J. O., & Wiberg, M. (2019). TestGardener: A program for optimal scoring and graphical analysis. In M. Wiberg, S. Culpepper, R. Janssen, J. González, & D. D. Molenaar (Eds.), *Quantitative psychology: 83rd annual meeting of the psychometric society* (pp. 87–94). Springer.

Li, Y., Cao, F., Cao, D., & Liu, J. (2015). Nursing students' post-traumatic growth, emotional intelligence and psychological resilience. *Journal of Psychiatric and Mental Health Nursing*, *22*, 326–332. https://doi.org/10.1111/jpm.12192

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review*, *8*, 290–300.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, *3*, 97–105. https://doi.org/10.1037/1528-3542.3.1.97

Ministry of Education and Culture. (2021). *Education system in Finland*. https://okm.fi/documents/1410845/15514014/Education+system+in+Finland/7c5a920b-47a5-c3ce-cbca-818ff3a5f848/Education+system+in+Finland.pdf?t=1631015290321

Orak, R. J., Farahani, M. A., Kelishami, F. G., Seyedfatemi, N., Banihashemi, S., & Havaei, F. (2016). Investigating the effect of emotional intelligence education on baccalaureate nursing Students' emotional intelligence scores. *Nurse Education in Practice*, *20*, 64–69. https://doi.org/10.1016/j.nepr.2016.05.007

Petrides, K., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology*, *98*, 273–289.

Pienimaa, A., Haavisto, E., & Talman, K. (2022). Emotional intelligence instruments used in healthcare education. *Journal of Nursing Education*, *61*(1), 6–11. https://doi.org/10.3928/01484834-20211130-01

Pienimaa, A., Talman, K., & Haavisto, E. (2021). The assessment of emotional intelligence in social care and healthcare student selection: A qualitative descriptive study. *Educational Research*, *63*(3), 302–318. https://doi.org/10.1080/00131881.2021.1936111

Polit, D., & Beck, C. (2006). The Content Validity Index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*, 489–497.

Ramsay, J. O., Li, J., & Wiberg, M. (2020). Better test scores with TestGardener. https://www.psych.mcgill.ca/misc/fda/downloads/FDAfuns/OptimalScoreBook.pdf

Rankin, B. (2013). Emotional intelligence: Enhancing values-based practice and compassionate care in nursing. *Journal of Advanced Nursing*, *69*(12), 2717–2725. https://doi.org/10.1111/jan.12161

Salminen-Tuomaala, M. H. (2020). Developing emotional intelligence and situational awareness through simulation coaching. *Clinical Nursing Studies*, *8*(2), 13–20. https://doi.org/10.5430/cns.v8n2p13

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, *46*(6), 401–435. https://doi.org/10.1080/00207594.2011.626049

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*, 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Shanta, L., & Gargiulo, L. (2014). A study of the influence of nursing education on development of emotional intelligence. *Journal of Professional Nursing*, *30*(6), 511–520. https://doi.org/10.1016/j.profnurs.2014.06.005

Snowden, A., Stenhouse, R., Young, J., Carver, H., Carver, F., & Brown, N. (2015). The relationship between emotional intelligence, previous caring experience and mindfulness in student nurses and midwives: A cross sectional analysis. *Nurse Education Today*, *35*(1), 152–158. https://doi.org/10.1016/j.nedt.2014.09.004

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. *Nurse Education Today 85*, 104271. https://doi.org/10.1016/j.nedt.2019.104271

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. https://doi.org/10.1891/1061-3749.22.1.94

The World Health Organisation (WHO). (2021). *The 2021 update, global health workforce statistics*. https://www.who.int/data/gho/data/themes/topics/health-workforce

Vierula, J., Karihtala T., Haavisto, E., & Talman, K. (2021). *Tutkimusraportti. Ammattikorkeakoulujen uuden digitaalisen valintakokeen kehittäminen – Toteuttamisvaiheen tulokset*. Metropolia Ammattikorkeakoulun julkaisuja TAITO-sarja 79. Metropolia ammattikorkeakoulu. https://urn.fi/URN:ISBN:978-952-328-288-9 (abstract available in English).

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*, *5*, 2549–2560. https://doi.org/10.1111/jan.14799

Wellhagen, G. J., Ueckert, S., Kjellsson, M. C., & Karlsson, M. O. (2021). An item response theory–informed strategy to model total score data from composite scales. *The AAPS Journal*, *23*(3), 45. https://doi.org/10.1208/s12248-021-00555-3

Zhang, M., & Cross, S. E. (2011). Emotions in memories of success and failure: A cultural perspective. *Emotion*, *11*(4), 866–880. https://doi.org/10.1037/a0024025

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pienimaa, A., Talman, K., Vierula, J., Laakkonen, E., & Haavisto, E. (2022). Development and psychometric evaluation of the Emotional Intelligence Test (EMI-T) for social care and healthcare student selection. *Journal of Advanced Nursing*, *00*, 1–14. https://doi.org/10.1111/jan.15557