Development and Psychometric Evaluation of the Telenursing Interaction and Satisfaction Scale (TISS)

Marie Mattisson (✉ marie.mattisson@liu.se)
  Linköping University: Linkopings universitet   https://orcid.org/0000-0001-5033-0874
Sussanne Börjeson
  Linkoping University Faculty of Medicine: Linkopings universitet Medicinska Fakulteten
Malou Lindberg
  Region Östergötland: Region Ostergotland
Kristofer Årestedt
  Linnaeus University: Linneuniversitet

Research

Abstract

Background: In telenursing, interaction between caller and telenurse is crucial for outcomes such as adherence, safety and satisfaction. There is a recurring demand for improved interaction in telenursing and a lack of measurement scales focusing on caller satisfaction with interaction. The purpose of this study was to develop and evaluate psychometric properties of the Telenursing Interaction and Satisfaction Scale (TISS).

Methods: This instrumental development stud was based on cross-sectional data. Callers to the National Medical Advisory Service in Sweden (n=616) completed a 60-item questionnaire, the Telenursing Interaction and Satisfaction Questionnaire (TISQ). Twenty-five of these items were selected to form the TISS. Selected items represent four dimensions of interaction according to the Interaction Model of Client Health Behavior; health information, professional-technical competence, affective support and decisional control. Data quality was evaluated in terms of missing data patterns and score distributions. Factor structure of the scale was evaluated with confirmatory factor analysis, convergent validity with Spearman correlations, internal consistency with ordinal alpha, scale reliability with composite reliability coefficients, and test-retest reliability with intraclass correlations.

Results: The amount of missing data was acceptable and equally distributed. The completeness of data was the highest for the subscale of professional-technical competencies (94%) and the lowest for the TISS total scale (80%). Data deviated significantly from a normal distribution, but all response options were endorsed. The CFA confirmed the hypothesized four-factor structure. Factor loadings ranged from 0.56 to 0.97, and factor correlations were high (0.88-0.96). Internal consistency (ordinal alpha = 0.82-0.97), scale reliability (0.88-0.99) and test-retest reliability (ICC = 0.77-0.86) were satisfactory for all scales.

Conclusion: The TISS holds satisfactory psychometric properties in the study sample. Findings support the use of four sub-scales for measuring caller satisfaction with interaction in telenursing. A total-score can be calculated and used in situations where multi-collinearity is a problem.

Background

The development and use of information and communication technology (ICT) within healthcare is growing, and even exploding as a result of the global Corona pandemic [1, 2]. Digital encounters provide sustainable and safe health care but also require communication competence that differs from face-to-face meetings [3]. In telenursing, telenurses triage a wide range of health conditions and provide verbal support, guidance and advice to entire populations [4]. The interaction between telenurse and caller is a complex and for telenurses cognitively demanding activity [5] depending on both caller and telenurse contribution [6]. According to Evans [7], exchange based on mutuality must be present for true interaction to occur and requirements on communication skills and interpersonal competence of telenurses are high.

There is a recurring demand for improved interaction in telenursing [8–10] and therefore also a need for evaluation instruments. These instruments should be developed and validated in a telenursing context and focus on interaction between the caller and the telenurse. Historically, few RCTs on nursing communication have been performed, partly due to the lack of standardized outcome measurements [11] and it is suggested that telenurses´ communication competence could be monitored by measures of satisfaction [12].

Satisfaction is a desired outcome of care, partly due to its importance for adherence [9, 13]. The concept of patient satisfaction has no widely adopted definition [14] but is thought of as a person-centered outcome that includes the evaluation of actual care received in relation to needs and expectations [15]. Factors contributing to satisfaction in healthcare include a wide set of variables that could be related either to the patient or to the healthcare provider, such as demographic variables, health status, expectations, organizational characteristics, interpersonal care, availability and outcome of care [16]. Among determinants to satisfaction with care, patients´ perceptions of interpersonal care is the most essential [16].

As a consequence, the Telenursing Interaction and Satisfaction Questionnaire (TISQ) was developed by Mattisson et al. [17]. The purpose of the TISQ is to explore caller satisfaction in telenursing and it thus includes a large set of determinants and confounders. All items are anchored and structured in accordance with a theoretical nursing model, the Interaction Model of Client Health Behavior (IMCHB), by Cox [18]. In this model, the interaction between client and healthcare professional is defined by four dimensions; health information, professional-technical competencies, affective support, and decisional control. The nurse should strive to tailor the interaction according to the needs of the client in order to attain outcomes such as satisfaction, improved health status and well-being [18]. The TISQ consists of a comprehensive set of 60 items. Twenty-five of these focus specifically on the interaction between caller and telenurse. The remaining 35 items cover other variables relevant for satisfaction in telenursing. An initial validation of the TISQ showed that all items had good content validity in a Swedish telenursing context and acceptable test-retest reliability on the item level [17].

The 25 items about interaction in the TISQ were developed as single items reflecting specific and detailed parts of the interaction from a caller perspective. However, single items have some drawbacks compared to multi-item scales. For example, the amount of random error is

commonly larger in single-item scores than in multi-item scale scores [19]. Also, the use of multi-item scales reduces the amount of data without substantial losses in precision and scope [20]. The 25 items could, if psychometrically sound, serve as a measuring scale.

Methods

Aim

The aim of this study was to develop and evaluate psychometric properties of the Telenursing Interaction and Satisfaction Scale (TISS), with a focus on data quality, factor structure, convergent validity and reliability.

Design

This instrument development study was based on cross-sectional data.

Description of Items

All 25 items in the TISS were selected from the TISQ [17] because of their focus on the interaction between the caller and the telenurse from a caller perspective. They represent four dimensions of interaction presented in a theoretical nursing model; the Interaction Model of Client Health Behavior (IMCHB) by Cox [18]; health information, professional-technical competencies, affective support, and decisional control. Consequently, items related to each of these dimensions were expected to measure the latent construct in four subscales. The subscale of health information includes eight items concerning the amount and relevance of information exchanged during an interaction. Professional-technical competencies include five items that measure the caller's opinion of the telenurse's medical knowledge and communication skills. Affective support includes nine items on how and to what degree the caller perceived the telenurse to be affectively supportive in terms of active listening, understanding, and empathy. The fourth subscale, decisional control, includes three items concerning the caller's experience of collaboration with the telenurse [17]. Two different response scales are used to rate the 25 items. Twenty of the items are rated on a four-point scale ranging from "Yes, completely" (1) to "No, not at all" (4). The remaining five items are rated on a five-point scale ranging from "Very satisfied" (1) to "Very dissatisfied" (5).

Data collection

Data were taken from a previous study that describes the development and validation of the TISQ [17]. In that study, data collection took place during five weeks in September-October of 2017. Callers were recruited when they called the Swedish National Medical Advisory Service, 1177. Each caller was automatically informed about the study and invited to participate on a voluntarily basis prior to their conversation with a telenurse. The message included information about the inclusion criteria: 18 years or older, health advisory call and not an administrative inquiry, and a health problem that directly pertained to the caller. Acceptance was communicated by pressing a certain dial code that produced a register of callers willing to participate in the study. A paper version of the Swedish 60-item TISQ, information about the study, and a form for written consent was sent by post within seven days of the call to 1400 callers who had volunteered to participate. In total, 622 of the 1400 questionnaires were returned (response rate 44.4%). Of the returned questionnaires, six had missing data on all 25 items about interaction and were excluded, leaving a sample of 616 callers for the evaluation of the TISS. Of those who completed the first questionnaire, 168 callers agreed to complete the questionnaire once more within two weeks for test-retest purposes. They received a second questionnaire and 114 (67.9%) of these questionnaires were returned.

Participants

A majority of the 616 callers were female (n = 424, 69%). The median age was 54.5 years (q1−q3 = 35-69.5). Most of the callers had prior experience of this specific advisory service (n = 537, 87%), and almost half of them (n = 251, 41%) reported the highest level of urgency prior to the call. After the call, about one third (n = 179, 29%) reported self-care advice as the main result of the call (Table 1).

Table 1

Descriptive data of callers (n = 616).

Age, median (q1–q3)	54.5 years (35–69.5)
Gender, n(%)	
Male	182 (29.5)
Female	424 (68.8)
None of the above	1 (0.2)
Educational level, n(%)	
Elementary school or less	90 (14.6)
Upper secondary school	213 (34.6)
University	299 (48.5)
Number of previous contacts with this service, n(%)	
None	68 (11.0)
One to five times	321 (52.1)
Six times or more	216 (35.1)
Caller's perceived level of urgency when calling, n(%)	
Not urgent at all	33 (5.4)
A little urgent	98 (15.9)
Urgent	229 (37.2)
Very urgent	251 (40.7)
Self-reported main result of the call, n(%)	
Self-care advice from the telenurse	179 (29.1)
Appointment with GP on call	90 (14.6)
Advice to visit ER	139 (22.6)
Advice to contact another caregiver	136 (22.1)
Other	59 (9.6)

Data Analysis

Descriptive statistics were used to present sample characteristics among callers and data quality. The 25 items in the TISS were treated as ordinal data in the psychometrical analyses.

Data quality was evaluated in terms of missing data patterns and score distributions for both item and scale scores. The frequencies of responses and score distributions for each item were examined to identify problems with missing data and floor and ceiling effects. Floor and/or ceiling effects were considered if the distribution exceeded 20% for the lowest or highest possible scores respectively [21]. The score distribution of the scales was evaluated using skewness and kurtosis statistics, the Kolmogorov-Smirnov (K-S) test, and plots (normal probability plots, normal quantile plots, and histograms). A normal distribution has skewness and kurtosis values close to 0 and a non-significant K-S test [22]. Completeness of data for the scale scores was also evaluated, i.e., the share of scores that could be computed without any imputation procedures.

The hypothesized four-factor structure of the 25 items in the TISS was evaluated using confirmatory factor analysis (CFA) with a weighted least square mean variance (WLSMV) estimation and polychoric correlations. Two CFA models were evaluated, one four-factor model and one higher-order factor model. Missing data were handled with pairwise deletion.

Model evaluation was assessed with chi-square goodness-of-fit statistics, Root Mean Square Error of Approximation (RMSEA) with 90% confidence interval, Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). Satisfactory

model fit was defined as: RMSEA close to 0.06 or below, CFI and TLI close to 0.95 or greater, and SRMR close to 0.08 or below [23].

Convergent validity was evaluated with Spearman's rho correlations between the TISS scale scores and one ordinal item on overall satisfaction with the call that was not included in the TISS. According to Cohen's conventions [24], correlations corresponding to a strong effect size ($r_s \geq 0.5$) were hypothesized since satisfaction with interaction is considered the most influential determinant to patient satisfaction [16].

Internal consistency was evaluated using ordinal alpha, an alternative to the traditional Cronbach's alpha for ordered categories [25]. An alpha level of 0.7 or higher is commonly described as satisfactory [26]. Scale reliability was evaluated using the composite reliability coefficient (CR), calculated according to Raykov [27] as illustrated by the equation below:

$$CR = \frac{\left(\Sigma \lambda_j \right)^2}{\left(\Sigma \lambda_j \right)^2 + \Sigma \epsilon_j}$$

Test-retest reliability was evaluated using the intraclass correlation coefficient (ICC) based on the sum of scores for the four subscales separately, and for the total scale score of all 25 items. The ICC analyses were based on a two-way mixed-effects model, absolute agreement. The ICC was interpreted in accordance with Koo and Li [28]: 0.5 poor, 0.5–0.75 moderate, 0.75–0.9 good, and > 0.9 excellent test-retest reliability.

The significance level was set at $p < 0.05$. The CFA analyses were conducted in Mplus 8.3 (Muthén and Muthén, Los Angeles, CA, USA), while internal consistency (ordinal alpha) was conducted in R statistical package 3.6.1 (the R Foundation for Statistical Computing, Vienna, Austria), including the psych package. All other analyses were conducted in SPSS statistics 24 (IBM Corp., Armonk, NY, USA).

Results

Data Quality

The frequency of missing data ranged from 0.5–6.3%, except for item 19 about opportunities to discuss alternative solutions to the health problem (10.9%) (Table 2). This implies a mean frequency of 4.6%. Completeness of data was the highest for the subscale of professional-technical competencies (n = 579, 94.0%), followed by health information (n = 567, 92.0%), affective support (n = 567, 92.0%), decisional control (n = 527, 85.6%), and TISS total scale (n = 495, 80.3%).

Table 2

Items and item statistics (n = 616).

Subscale according to the IMCHB[†]	Item in TISS[‡]	Item in TISQ[§]	Wording of item	Missing data, n (%)	Score distribution, n (%)			
			"When you called, did you perceive that…"		Yes, completely (1)	Yes, partially (2)	No, not quite (3)	No, not at all (4)
Health information	1	B1a	…you were given the opportunity to ask all your questions?	24 (3.9)	500 (81.2)	73 (11.9)	14 (2.3)	5 (0.8)
	2	B1b	…you received answers to all your current questions at the time?	27 (4.4)	374 (60.7)	159 (25.8)	40 (6.5)	16 (2.6)
	3	B1c	…the nurse provided you with information on the future potential development of the health problem?	39 (6.3)	185 (30.0)	181 (29.4)	106 (17.2)	105 (17.0)
	4	B1d	…you got information about what you should do next?	25 (4.1)	423 (68.7)	124 (20.1)	27 (4.4)	17 (2.8)
	5	B1e	…you had understood the advice/information when ending the call?	34 (5.5)	509 (82.6)	57 (9.3)	8 (1.3)	8 (1.3)
	6	B1f	…you received advice and information adapted to your needs and conditions at the time?	30 (4.9)	441 (71.6)	97 (15.7)	22 (3.6)	26 (4.2)
	7	B1g	…you were informed about where to find additional information?	37 (6.0)	265 (43.0)	142 (23.1)	75 (12.2)	97 (15.7)
Professional-technical competencies	8	B3a	…the nurse had enough competence to deal with your health problem?	16 (2.6)	381 (61.9)	157 (25.5)	45 (7.3)	17 (2.8)
	9	B3b	…the nurse asked relevant questions about your health problem?	29 (4.7)	417 (67.7)	134 (21.8)	19 (3.1)	17 (2.8)
	10	B3c	…the nurse was thorough in his/her work?	32 (5.2)	447 (72.6)	103 (16.7)	24 (3.9)	10 (1.6)
	11	B3d	…the nurse was skilled in leading the conversation forward?	29 (4.7)	416 (67.5)	119 (19.3)	33 (5.4)	19 (3.1)
Affective support	12	B5a	…you felt confidence in the nurse you talked to?	23 (3.7)	441 (71.6)	104 (16.9)	35 (5.7)	13 (2.1)
	13	B5b	…the nurse listened	32 (5.2)	475 (77.1)	85	15 (2.4)	9 (1.5)

	#	Code	Question					
			attentively?		(13.8)			
	14	B5c	…the nurse understood what you wanted?	25 (4.1)	469 (76.1)	99 (16.1)	16 (2.6)	7 (1.1)
	15	B5d	…the nurse showed empathy?	38 (6.2)	402 (65.3)	128 (20.8)	28 (4.5)	20 (3.2)
	16	B5e	…the nurse was friendly?	32 (5.2)	517 (83.9)	47 (7.6)	13 (2.1)	7 (1.1)
	17	B5f	…the nurse was calm and instilled a sense of security?	35 (5.7)	498 (80.8)	56 (9.1)	20 (3.2)	7 (1.1)
	18	B5g	…the nurse showed an interest in your understanding of the health problem?	38 (6.2)	421 (68.3)	102 (16.6)	37 (6.0)	18 (2.9)
Decisional control	19	B10a	…you were given opportunities to discuss alternative solutions to the health problem?	67 (10.9)	199 (32.3)	176 (28.6)	97 (15.7)	77 (12.5)
	20	B10b	…you and the nurse agreed on how to deal with your health problem?	36 (5.8)	427 (69.3)	99 (16.1)	35 (5.7)	19 (3.1)

			"Overall, how satisfied were you with…"	Missing data, n (%)	Very satisfied (1)	Satisfied (2)	Neither satisfied nor dissatisfied (3)	Dissatisfied (4)	Very dissatisfied (5)
Health information	21	B2	…the advice and information you were given?	3 (0.5)	366 (59.4)	167 (27.1)	48 (7.8)	16 (2.6)	16 (2.6)
Prof/technical competencies	22	B4	…the competence of the nurse?	11 (1.8)	380 (61.7)	147 (23.9)	52 (8.4)	18 (2.9)	8 (1.3)
Affective support	23	B6	…the nurse's ability to support you affectively?	11 (1.8)	366 (59.4)	143 (23.2)	65 (10.6)	20 (3.2)	11 (1.8)
	24	B7	…how the nurse treated you?	20 (3.2)	424 (68.8)	109 (17.7)	39 (6.3)	15 (2.4)	9 (1.5)
Decisional control	25	B11	… the possibility to influence the result of the call?	14 (2.3)	324 (52.6)	167 (27.1)	80 (13.0)	17 (2.8)	14 (2.3)

[†] Interaction Model of Client Health Behavior (Cox, 1982)

[‡] Telenursing Interaction and Satisfaction Scale

[§] Telenursing Interaction and Satisfaction Questionnaire (Mattisson et al., 2019)

All items displayed a positive score distribution with floor effects (i.e. high satisfaction rates). Data deviated significantly from a normal distribution. Normality tests for scales were as follows: health information (skewness = 1.59, K-S(567) = 0.162, $p < 0.001$), professional-technical competencies (skewness = 1.87, K-S(579) = 0.250, $p < 0.001$), affective support (skewness = 2.47, K-S(567) = 0.277, $p < 0.001$), decisional control (skewness = 1.00, K-S(527) = 0.177, $p < 0.001$), and the TISS total scale (skewness = 1.98, K-S(495) = 0.191, $p < 0.001$). The findings were supported with an examination of the distribution plots. All response options were endorsed (Table 2).

Factor Structure

The four-factor model demonstrated good model fit according to all fit indices except the chi-square goodness-of-fit: RMSEA = 0.059, CFI = 0.988, TLI = 0.987, SRMR = 0.039. The factor loadings were all significant at a level of $p < 0.001$ and ranged from 0.56 to 0.94 for health information, 0.83 to 0.97 for professional-technical competencies, 0.86 to 0.97 for affective support, and 0.66 to 0.95 for decisional control (Table 3). The factor correlations between the four subscales ranged from 0.88 to 0.96 (Table 4).

The higher order four-factor model demonstrated good model fit according to the fit indices: RMSEA = 0.061, CFI = 0.987, TLI = 0.986, and SRMR = 0.041. The factor loadings between the separate items in the TISS and the four latent factors were the same as for the four-factor model (Table 3). The factor loadings between the four latent factors and the higher order factor were 0.96 for health information, 0.99 for professional-technical competencies, 0.94 for affective support, and 0.94 for decisional control, all significant at a level of $p < 0.001$ (Table 4).

Table 3
Standardized factor loadings, residual variances, and goodness-of-fit indices for the four-factor model and the higher-order model.

Item/subscale in the TISS	Four-factor CFA model		Higher-order CFA model	
	Factor loading (λ)	Residual variance (δ)	Factor loading (λ)	Residual variance (δ)
Health information				
1	0.86	0.25	0.86	0.26
2	0.85	0.28	0.85	0.28
3	0.65	0.57	0.65	0.57
4	0.81	0.34	0.81	0.34
5	0.76	0.42	0.76	0.42
6	0.91	0.17	0.91	0.17
7	0.56	0.69	0.56	0.69
21	0.94	0.11	0.94	0.11
Professional-technical competencies				
8	0.83	0.31	0.83	0.32
9	0.90	0.19	0.90	0.19
10	0.95	0.09	0.95	0.09
11	0.95	0.10	0.95	0.10
22	0.97	0.06	0.97	0.06
Affective support				
12	0.94	0.11	0.94	0.11
13	0.94	0.12	0.94	0.12
14	0.86	0.26	0.85	0.26
15	0.90	0.18	0.90	0.18
16	0.93	0.13	0.93	0.13
17	0.92	0.14	0.92	0.14
18	0.91	0.16	0.91	0.16
23	0.94	0.11	0.94	0.11
24	0.97	0.06	0.97	0.06
Decisional control				
19	0.66	0.56	0.66	0.56
20	0.90	0.19	0.90	0.19
25	0.95	0.10	0.95	0.10
Goodness-of-fit values[†]				
Chi-square (df)	843.0 (269)		893.1 (271)	
Chi-square p-value	< 0.001		< 0.001	
RMSEA (90% CI) [LL, UL]	0.059 [0.054, 0.063]		0.061 [0.057, 0.066]	
RMSEA p-value	0.001		< 0.001	

	Four-factor CFA model	Higher-order CFA model
CFI	0.988	0.987
TLI	0.987	0.986
SRMR	0.039	0.041

[†] Criteria for satisfactory model fit (Hu & Bentler, 1999):

Chi-square goodness-of-fit: non-significant

RMSEA (Root Mean Square Error of Approximation): close to 0.06 or below

CFI (Comparative Fit Index): close to 0.95 or greater

TLI (Tucker-Lewis Index): close to 0.95 or greater

SRMR (Standardized Root Mean Square Residual): close to 0.08 or below.

Table 4
Factor correlations and higher-order factor loadings.

	Four-factor model correlations				Higher-order factor loadings
Subscales in the TISS	HI	PTC	AS	DC	
Health information (HI)					0.96
Professional-technical competencies (PTC)	0.96				0.99
Affective support (AS)	0.89	0.94			0.94
Decisional control (DC)	0.93	0.88	0.90	-	0.94
All factor correlations and factor loadings are significant at a level of p < 0.001.					
TISS = Telenursing Interaction and Satisfaction Scale					

Convergent validity

Correlations between the TISS scale and the item on overall satisfaction were all above the expected value ($r_s \geq 0.5$): 0.64 for health information, 0.71 for professional technical competencies, 0.69 for affective support and 0.63 for decisional control. The strongest correlation was found between the TISS total and overall satisfaction ($r_s = 0.74$).

Internal Consistency

The ordinal alpha was above the recommended level of 0.7 for all scales with a range between 0.82 for the subscale of decisional control and 0.97 for the TISS total scale (Table 5).

Scale Reliability

The scale reliability of subscales varied from 0.88 to 0.98. It was lowest for the subscale of decisional control and the highest for the subscale of affective support. Scale reliability for the TISS total scale was 0.99 (Table 5).

Test-retest Reliability

The ICC ranged from 0.77 to 0.86, indicating good test-retest reliability for all subscales and the TISS total scale. However, the lower end of the confidence interval overlapped the lower limit for good stability regarding the subscales of affective support and decisional control (Table 5).

Table 5
Internal consistency, scale reliability and test-retest reliability (n = 114).

TISS scales	Ordinal alpha	Composite reliability	ICC	95% CI for ICC
Health information (8 items)	0.92	0.93	0.84	0.77, 0.89
Professional-technical competencies (5 items)	0.95	0.96	0.86	0.79, 0.90
Affective support (9 items)	0.96	0.98	0.77	0.68, 0.84
Decisional control (3 items)	0.82	0.88	0.77	0.67, 0.84
TISS total scale (25 items)	0.97	0.99	0.86	0.80, 0.91
All ICCs significant at a level of p < 0.001				
TISS = Telenursing Interaction and Satisfaction Scale				

Discussion

This study presents a psychometric evaluation of the Telenursing Interaction and Satisfaction Scale (TISS) intended to measure satisfaction with the interaction that takes place between telenurses and callers from the perspective of the caller. The analysis of data quality, factor structure, convergent validity, and reliability in a sample of 616 callers showed that the TISS holds satisfactory psychometric properties. To the best of our knowledge, this makes the TISS the first valid and reliable scale to measure caller satisfaction with interaction in telenursing.

The overall amount of missing data was acceptable and equally distributed among items with one exception; item 19, concerning opportunities to discuss alternative solutions to the health problem with the telenurse. The high frequency of missing data in this item might be explained by the callers' perception that no relevant alternative existed. The item was judged as highly relevant in an earlier study [17] and adding a response option, "Not applicable," might be appropriate in future refinement of the TISS. Despite this, the TISS demonstrated satisfactory data quality in terms of completeness of data, which was acceptable for all sub-scales.

All scales in the TISS produced non-normally distributed data with high degrees of satisfaction which is in line with other satisfaction studies in telenursing [29]. Skewness could result in problems with sensitivity and responsiveness of measuring scales [20]. The skewed result in this study could have several explanations such as non-response bias [30] or low expectations among callers resulting in high satisfaction rates [31]. Reasons for skewness can also be found in the scale construction. According to Voutilainen et al. [32], items in a satisfaction questionnaire should be neutrally worded and sufficiently many to "increase the likelihood that the least satisfactory care components are also included" [32]. This corresponds well with the 25 neutrally worded items in the TISS. Even though data in this study were skewed, every response option in the TISS was used indicating that the response scales are relevant. This is further supported by the result of previous cognitive interviews with callers who found all items and response scales relevant and comprehensible [17].

The hypothesized factor structure was confirmed by all fit indices except for the chi-square goodness-of-fit statistics. It is well known that the chi-square test is highly sensitive to large samples, and it is therefore commonly used for purposes other than examining model fit, e.g. to compare nested CFA models [33]. Since all other goodness-of-fit indices indicated good to excellent model fit, the poor chi-square test can probably be explained by the large sample in this study. Consequently, the results support the TISS as a multidimensional scale that measures satisfaction with the four theoretical dimensions of interaction. A higher-order factor model was evaluated due to strong factor correlations. The goodness-of-fit indices of this latter model were somewhat impaired compared to the four-factor model, but still acceptable. It could be argued that correlations between the four factors were too high, and that the subscales could be questioned. Correlations of 0.85 according to Brown [33] are considered an acceptable upper limit of factor correlations, and models with higher values should be revised in order to reduce multi-collinearity between factors. However, Brown [33] also advises against re-specification without a theoretical rationale. In this study, the factors were theoretically driven, and there was no clear rationale for re-specification. Therefore, we recommend using the subscale scores instead of the total score. The total scale score can be used in situations where it can be a problem to use the four subscales, for example problems with multi-collinearity in regression analysis.

Internal consistency (ordinal alpha) was high for both the TISS subscales and the total scale, supporting the homogeneity among items. The ordinal alpha increases with increasing number of items, and according to Streiner et al. [20], high alpha values can be seen as an indication of redundancy among items. This was taken into account in the content validation of the TISQ, where eight items on interaction and with acceptable I-CVIs were deleted due to comments on redundancy [17]. The remaining items were considered to measure separate aspects of interaction with the telenurse. Another indication of redundancy is correlated residuals [33], a problem that was not identified in this study. Thus, our findings support the hypothesis that the four dimensions of interaction have high content validity and internal consistency at the same time, as discussed by Keszei et al. [19].

Both scale and test-retest reliability were satisfactorily high. For two of the subscales, affective support and decisional control, the lower limit of the 95% CI of the composite reliability coefficient fell right below the critical value of 0.7 and may require further evaluation in future studies.

The results from this study, in addition to previous findings about the content validity of included items [17], support the use of the TISS as an outcome measure in clinical and research settings where the interaction between a caller and a telenurse is in focus. It was developed and evaluated in a context where telenurses provide nursing care without visual input and should therefore primarily be used in situations where visual input is not at hand. The profound relation between interaction and safety in telenursing [34–36] could make the TISS a useful instrument in patient safety work. It also responds well to individual telenurses' call for individual feedback on work performance, as described by Wahlberg & Bjorkman [37], or for comparisons between different telenursing sites or subgroups in a population.

The TISS scale scores should be calculated in three steps: First, all item scores need to be reverse coded so that high scores reflect high levels of satisfaction. Second, sub-scale scores are calculated by summating all item responses within each sub-scale. This implies that the possible sub-scale score range is 8–33 for health information, 5–21 for professional technical competence, 9–38 for affective support, and 3–13 for decisional control. Third, since the sub-scales include different numbers of items, the scores should be linearly transformed into a 0–10 scale using the following formula: ((raw scale score − lowest possible score) / possible score range) × 10. The TISS total scale score range is 25–105 and can be linearly transformed in the same manner as the sub-scales.

Methodological Considerations

The study has some limitations that need to be considered. First, no power calculation was conducted. However, the sample size of 616 callers is more than sufficient according to the rule-of-thumb for five or ten observations per estimated parameter. Additionally, WLSMV is less restrictive than other estimators for categorical indicator variables [33]. The study also has some important strengths. The ordinal nature of data was considered in all statistical analyses in order to avoid underestimating covariances in the CFA [33]. Another strength is that every caller who contacted the service during the inclusion period was invited to participate in the study. The sample showed satisfactory variation in age, sex, and other demographic variables. However, this variation is not a high priority in psychometric evaluation studies where sufficient sample size is considered more important [33]. Further psychometric evaluation could include aspects of responsiveness and sensitivity [20], especially if change over time is the prime objective. Before use in other countries, the TISS must be properly translated and validated [38].

Conclusion

The TISS offers a possibility to measure satisfaction with interaction in telenursing, with satisfactory validity and reliability. The four sub-scales in the TISS (health information, professional-technical competence, affective support and decisional control) are anchored in nursing theory, and strengthens its theoretical underpinnings. Findings support the use of all four sub-scales. A total-score can be calculated and used in situations where multi-collinearity is a problem. The TISS can be used for both clinical and research purposes in a telenursing context where caller satisfaction with interaction is in focus.

Abbreviations

CFA: Confirmatory factor analysis

CFI: Comparative Fit Index

CI: Confidence interval

CR: Composite reliability coefficient

ICC: Intraclass correlation coefficient

ICT: Information and communication technology

I-CVI: Item-Content Validity Index

IMCHB: Interaction Model of Client health behavior

RMSEA: Root mean square error of approximation

SRMR: Standardized root mean square residual

TISQ: Telenursing interaction and satisfaction questionnaire

TISS: Telenursing interaction and satisfaction scale

TLI: Tucker-Lewis index

WLSMV: Weighted least square mean variance

Declarations

References

1. Kannampallil T, Ma J. Digital Translucence: Adapting Telemedicine Delivery Post-COVID-19. Telemedicine and e-Health. 2020;26:1120-1122. DOI: 10.1089/tmj.2020.0158
2. Bokolo AJ. Application of telemedicine and eHealth technology for clinical services in response to COVID-19 pandemic. Health and Technology. 2021;DOI: 10.1007/s12553-020-00516-4
3. Roberts LC, Osborn-Jenkins L. Delivering remote consultations: Talking the talk. Musculoskeletal Science and Practice. 2020;DOI: 10.1016/j.msksp.2020.102275
4. Souza-Junior VD, Mendes IAC, Mazzo A, Godoy S. Application of telenursing in nursing practice: an integrative literature review. Applied Nursing Research. 2016;29:254-260. DOI: 10.1016/j.apnr.2015.05.005
5. Mataxen PA, Webb LD. Telehealth nursing: More than just a phone call. Nursing. 2019;49:11-13. DOI: 10.1097/01.NURSE.0000553272.16933.4b
6. Yliluoma P, Palonen M. Telenurses' experiences of interaction with patients and family members: nurse-caller interaction via telephone. Scandinavian journal of caring sciences. 2020;34:675-683. DOI: 10.1111/scs.12770
7. Evans EC. Exploring the Nuances of Nurse-Patient Interaction through Concept Analysis. Nursing Science Quarterly. 2016;29:62-70. DOI: 10.1177/0894318415614904
8. Röing M, Rosenqvist U, Holmström I. Threats to patient safety in telenursing as revealed in Swedish telenurses' reflections on their dialogues. Scandinavian Journal of Caring Sciences. 2013;27:969-976. DOI: 10.1111/scs.12016
9. Purc-Stephenson RJ, Thrasher C. Patient compliance with telephone triage recommendations: A meta-analytic review. Patient Education & Counseling. 2012;87:135-142.

10. Ernesäter A, Engström M, Winblad U, Rahmqvist M, Holmström I. Telephone nurses' communication and response to callers' concern—a mixed methods study. Applied Nursing Research. 2016;29:116-121. DOI: 10.1016/j.apnr.2015.04.012

11. Kerr D, Ostaszkiewicz J, Dunning T, Martin P. The effectiveness of training interventions on nurses' communication skills: A systematic review. Nurse Education Today. 2020;89:104405. DOI: https://doi.org/10.1016/j.nedt.2020.104405

12. Rahmqvist M, Ernesäter A, Holmström I. Triage and patient satisfaction among callers in Swedish computer-supported telephone advice nursing. Journal of Telemedicine & Telecare. 2011;17:397-402.

13. Williams B, Warren S, McKim R, Janzen W. Caller self-care decisions following teletriage advice. Journal of Clinical Nursing. 2012;21:1041-1050. DOI: 10.1111/j.1365-2702.2011.03986.x

14. Batbaatar E, Dorjdagva J, Luvsannyam A, Amenta P. Conceptualisation of patient satisfaction: a systematic narrative literature review. Perspectives in Public Health. 2015;135:243-250. DOI: 10.1177/1757913915594196

15. Larson E, Sharma J, Bohren MA, Tuncalp Ö. When the patient is the expert: measuring patient experience and satisfaction with care. Bulletin of the World Health Organization. 2019;97:563-569. DOI: 10.2471/BLT.18.225201

16. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM. Determinants of patient satisfaction: a systematic review. Perspectives in Public Health. 2017;137:89-101. DOI: 10.1177/1757913916634136

17. Mattisson M, Johnson C, Börjeson S, Årestedt K, Lindbert M. Development and content validation of the Telenursing Interaction and Satisfaction Questionnaire (TISQ). Health Expectations. 2019;22:1213-1222. DOI: 10.1111/hex.12945

18. Cox CL. An interaction model of client health behavior: theoretical prescription for nursing. Advances in Nursing Science. 1982;5:41-56.

19. Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. Journal of Psychosomatic Research. 2010;68:319-323. DOI: 10.1016/j.jpsychores.2010.01.006

20. Streiner, Norman GR, Cairney J. Health measurement scales : a practical guide to their development and use. Fifth edition. ed. Oxford: Oxford University Press; 2015.

21. Hagell P, Alvariza A, Westergren A, Årestedt K. Assessment of Burden Among Family Caregivers of People With Parkinson's Disease Using the Zarit Burden Interview. Journal of Pain and Symptom Management. 2017;53:272-278. DOI: https://doi.org/10.1016/j.jpainsymman.2016.09.007

22. Field AP. Discovering statistics using IBM SPSS statistics. 5th ed. London: Sage Publications; 2018.

23. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal. 1999;6:1-55. DOI: 10.1080/10705519909540118

24. Cohen J. A power primer. Psychological Bulletin. 1992;112:155-159. DOI: 10.1037/0033-2909.112.1.155

25. Gadermann AM, Guhn M, Zumbo BD. Estimating Ordinal Reliability for Likert-Type and Ordinal Item Response Data: A Conceptual, Empirical, and Practical Guide. Practical Assessment, Research & Evaluation. 2012;17:

26. Nunnally JC, Bernstein IH. Psychometric theory. New York: McGraw-Hill; 1994.

27. Raykov T. Estimation of Composite Reliability for Congeneric Measures. Applied Psychological Measurement. 1997;21:173-184. DOI: 10.1177/01466216970212006

28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine. 2016;15:155-163. DOI: 10.1016/j.jcm.2016.02.012

29. Lake R, Georgiou A, Li J, Li L, Byrne M, Robinson M, Westbrook JI. The quality, safety and governance of telephone triage and advice services - an overview of evidence from systematic reviews. BMC health services research. 2017;17:614. DOI: 10.1186/s12913-017-2564-x

30. Boscardin CK, Gonzales R. The Impact of Demographic Characteristics on Nonresponse in an Ambulatory Patient Satisfaction Survey. Joint Commission Journal on Quality & Patient Safety. 2013;39:123-128.

31. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. Social Science & Medicine. 1997;45:1829-1843.

32. Voutilainen A, Pitkäaho T, Vehviläinen-Julkunen K, Sherwood PR. Meta-analysis: methodological confounders in measuring patient satisfaction. Journal of Research in Nursing. 2015;20:698-714. DOI: 10.1177/1744987115619209

33. Brown TA. Confirmatory factor analysis for applied research. Second ed. New York: The Guilford Press; 2015.

34. Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to non-normative symptom description and interpersonal communication: a mixed methods study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2017;25:52. DOI: 10.1186/s13049-017-0390-0

35. Ernesäter A, Winblad U, Engström M, Holmström IK. Malpractice claims regarding calls to Swedish telephone advice nursing: what went wrong and why? Journal of Telemedicine & Telecare. 2012;18:379-383. DOI: 10.1258/jtt.2012.120416

36. Ericsson M, Ängerud KH, Brännström M, Lawesson SS, Strömberg A, Thylén I. Interaction between tele-nurses and callers with an evolving myocardial infarction: Consequences for level of directed care. European Journal of Cardiovascular Nursing. 2019;18:545-553. DOI:

10.1177/1474515119848195

37. Wahlberg AC, Bjorkman A. Expert in nursing care but sometimes disrespected—Telenurses' reflections on their work environment and nursing care. Journal of Clinical Nursing. 2018;27:4203-4211. DOI: 10.1111/jocn.14622

38. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. Journal of Evaluation in Clinical Practice. 2011;17:268-274. DOI: 10.1111/j.1365-2753.2010.01434.x