Global Nurs. Insights 2025,12(2), 100-108

Research Article

Turkish Validity and Reliability of the Systems Thinking Scale

Gizem AÇIKGÖZ[1]

ABSTRACT

Aim: The aim of the study is to evaluate the validity and reliability of the Turkish version of the System Thinking Scale.

Materiel and Methods: The methodological study was conducted with 369 health professionals working in Istanbul between February - July 2024. The data were collected with descriptive information form and the Systems Thinking Scale. SPSS 24.0 program was used to evaluate the data, and the validity and reliability tests were conducted.

Results: The Kaiser Meyer Olkin coefficient of the scale was found 0.87, and the scale was suitable for factor analysis. The factor loadings of the items varied between 0.50 and 0.89, and the total variance explained by the scale was 55.48%. The scale consisted of a single dimension. The fit indices of the scale were $\chi^2/df = 4.68$, RMSEA = 0.09, NFI = 0.90, NNFI = 0.90, SRMR = 0.08, and CFI = 0.95. The AVE was 0.535, CR was 0.957, ICC was 0.99, and α was 0.95.

Conclusions: The Turkish version of the system thinking scale consists 20 items and a single dimension. The scale, which demonstrated good psychometric properties in Turkish, can be used to evaluate systems thinking within health systems.

Implication for nursing practice/management or policy: It is thought that the use of the scale will create awareness for ensuring system security in health systems and contribute to the orientation of considering the system as a whole.

Keywords: Quality of Health, Reliability, System Thinking, Validity

[1] Assist. Prof., İstanbul Kent University, Faculty of Health Science, Nursing, İstanbul, Türkiye, E-mail: gzmacikgoz@gmail.com, Tel: 5066884109, ORCID: 0000-0002-6133-9038

Received: 14 November 2024, Accepted: 14 February 2025

Citation: Açıkgöz, G. (2025). Turkish Validity and Reliability of the Systems Thinking Scale. *Global Nursing Insights, 12*(2), 100-108.
DOI: 10.31125/globnursinsights.1743913

Global Nurs. Insights 2025,12(2), 100-108

INTRODUCTION

Problems that arise in individual or professional life are quite interrelated, but the cause-and-effect relationship that is often put forward is insufficient in solving problems (Dolansky et al., 2020; Plack et al., 2018). Systems thinking is a way of thinking that takes over all aspects of the current situation with a broader perspective in this process and produces more effective results. It is an approach that targets the relationships between the parts that make up the problem. This thought includes an understanding of how actions will affect each other, how they can strengthen each other or how they can neutralize each other (Phillips et al., 2018). Systems thinking is based on the General System Theory put forward by Ludwig von Bertalanffy. According to this theory, a system is a whole consisting of inseparable parts. This approach, which also forms the basis of systems thinking, adopts three approaches: holistic approach, interdisciplinary approach and scientific approach. Holistic approach considers the system that interacts with each other and the external environment. Interdisciplinary approach means that different scientific experts come together to develop different approaches to the problem, make decisions and produce solutions in line with these approaches. And the scientific approach refers to making sufficient use of observation, experiment and evidence-based methods in the solution process (Arnold and Wade, 2015; Tetuan et al, 2018). Therefore, system thinking is a roadmap used in the in-depth examination and analysis of events and situations with a scientific approach. It has great importance in the successful maintenance of managerial and organizational processes (Burmakoğlu et al., 2016; Plack et al., 2018). The health system is one of the most important systems that requires a holistic perspective due to its dynamic and complex structure and therefore requires the adoption of a system approach. System thinking has a key role in all managerial processes in the health system in terms of considering events, situations or formations from a broad perspective and focusing on the relationships between the parts of the whole. Quality management is one of the most important ones in particular (Batalden and Mohr, 1997; Burmakoğlu et al., 2018). System thinking is accepted as one of the six basic competencies for physicians by the Accreditation Council of Graduate Medical Education (ACGME) and as one of the basic competencies for nurses by the Future of Nursing Report (FNR) (Dolansky et al., 2020). International health education authorities state that systems thinking has an important place in health education and research. International organizations dealing with quality and patient safety in health such as the London Protocol, Systems Engineering Initiative for Patient Safety, Critical Incident Analysis and Systems Awareness Model also emphasize the importance of systems thinking (Batalden and Mohr, 1997; Dolansky et al., 2020). The World Health Organization (WHO) emphasizes the importance of the nursing profession for the active use and dissemination of systems thinking. It states that nurses, as a professional group that is in active communication with patients and other health professionals, play an important role in this regard (World Health Organisation, 2020).

Studies emphasize that systems thinking has an important place in the success of the health system, but show that this system of thinking remains mostly theoretical and cannot be sufficiently transferred to practice (Rusoja et al., 2018). So, it is thought that it would be useful to evaluate the status of systems thinking first. However, no Turkish measurement tool has been found in the literature that measures this system of thinking.

Aim

The aim of the study is to evaluate the validity and reliability of the Systems Thinking Scale developed in English and to adapt it to the Turkish language.

Hypotheses/Study Questions

Is the Turkish form of the Systems Thinking Scale a valid scale?
Is the Turkish form of the Systems Thinking Scale a reliable scale?
Is the Turkish form of the Systems Thinking Scale a suitable scale for use by health professionals?

MATERIAL and METHODS

Study Design

This study is a methodological study conducted to test the validity and reliability of the Systems Thinking Scale in Turkish language and culture.

Study Sample

The research was conducted with healthcare professionals working in healthcare institutions in Istanbul between February and July 2024. The snowball method was used to determine the sample of the research, and the research was conducted with healthcare professionals who volunteered to participate in the research. In scale development and adaptation studies, sample size is important for the validity and reliability of the scale (Hair et al., 1995) The sample size should be more than 100 people, while Tabachnick and Fidell (2018) emphasized that it should be at least 300 people (Baştürk et al., 2013; Tabachnick and Fideli, 2018; Yurdabakan and Çüm, 2017) . The research sample was targeted to be over 300 participants in the light of the literature, and the research was conducted with 369 participants.

Data Collection Tools

Descriptive Information Form and Systems Thinking Scale were used.

Descriptive Information Form: It consists of 8 questions prepared considering the literatüre (Bolarinwa, 2015; Çapık et al., 2018).

Systems Thinking Scale: The scale was developed by Dolansky and his colleagues in 2019. The scale is a 20-item, one-dimensional and 5-point Likert Type Scale (1- never, 2 -rarely, 3-sometimes, 4-often and 5-often). The scale includes no reverse question. The

total score of the scale varies between 20-100 and and it is interpreted that as the scale score increases, the systems thinking skill increases (Dolansky et al., 2020).

Data Collection

Data was collected using online data collection tools, and in the collection of data, health professionals from different professions were reached in accordance with the snowball method, and through these health professionals, the data collection tool was delivered to the participants a total of 6 times in a 6-month period. The data collection process was completed when the targeted sample size of over 300 participants was reached.

The research was carried out in 4 stages: linguistic validity, content validity, construct validity and reliability.

Linguistic Validity

Scale items were translated into Turkish by a researcher with English proficiency and checked by 2 English Language Experts. The Turkish translation of the scale was translated back into English by 2 different English Language Experts and the compatibility of the original and English translation versions was evaluated by 2 Language Experts. The scale was finalized with the changes made after the suggestions from the experts (Davis, 1992).

Content Validity

Davis Technique was used in the evaluation of the scope validity of the scale. The scale items were presented to 13 health professionals who are experts in their fields, and they were asked to evaluate them as “1-Not relevant to the subject”, “2-Much correction is required”, “3-Related to the subject but little correction is required” and “4-Related to the subject” in accordance with the Davis Technique. In the Davis technique, the item content validity ratio (CVR) was calculated by dividing the number of experts who marked the options “related to the subject” and “related to the subject but little correction is required” by the total number of experts who responded to the item, and the content validity index (CVI) was calculated by averaging the scope validity ratios of the items (Bolarinwa, 2015; Gökdemir and Yılmaz, 2023).

Construct Validity

Construct validity was evaluated using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), and the suitability of the scale for factor analysis was determined using Bartlett's Sphericity test and Keiser-Mayer-Olkin (KMO) tests. In addition, convergent-discriminant validity was tested to verify the results obtained with factor analysis (Gökdemir and Yılmaz, 2023; Stevens and Mahwah, 2002).

Reliability

The reliability of the scale over time was evaluated using the test-retest method, and the data collection tool was re-administered to 30 participants at a 4-week interval. The internal consistency of the scale was assessed using Cronbach's alpha (Bolarinwa, 2015; Korku, 2002).

Data Analysis

SPSS 24.0 program was used in the analysis of the data. Descriptive statistical methods (frequency, ratio, percentage) were used in the evaluation of the data. The content validity of the scale was evaluated by calculating CVR and CVI in accordance with the Davis Technique (Davis, 1992). Before the construct validity of the scale, its suitability for factor analysis was evaluated with Bartlett Sphericity test and KMO tests, and the factor loadings and total explained variances of the scale items, CFA fit indices ($\chi 2$ / df , RMSE, NFI, NNFI, SRMR and CFI) and the average explained variance (AVE) and composite reliability (CR) in convergent and discriminant validity were calculated with EFA (Bozkurt and Yılmaz, 2024; Karakoç and Dönmez, 2014; Korku, 2002) . For the reliability of the scale over time, t-test, Pearson correlation analysis and intraclass correlation (ICC) analysis were performed in paired groups. Cronbach's alpha value was calculated for the internal consistency of the scale (Zivadinovic, 2004).

Ethical Considerations

This study complies with the Helsinki Declaration of Principles and complies with Research and Publication Ethics. The ethical consideration of the study was assessed by the ethics committee of a foundation university in Istanbul. Before the data was collected from the participants in the study, the “Informed Voluntary Consent Form” prepared by the researcher was added to the first page of the online survey form, and the participants were allowed to fill out the survey after approving it. Permission was obtained from the scale owner via e-mail to verify the validity and reliability of the scale in Turkish.

Limitations

The main limitation of the study is that the participants' descriptive characteristics such as age, gender, marital status, education level, profession, type of institution they work in, and unit they work in could not be homogenized. The study was conducted with health professionals working in Istanbul to provide homogeneous data, and the study data is limited to Istanbul.

RESULTS

It was found that 28.2% (n=104) of the participants were 40 years old and above, 57.5% (n=212) were female, 50.9% (n=188) were married, 38.5% (n=142) had a bachelor's degree, 44.4% (n=164) were nurses, 56.4% (n=208) of them worked in a public hospital and 42.8% (n=158) of them work in a in-patient service (Table 1).

Global Nurs. Insights 2025,12(2), 100-108

Table 1. Socio-demographic Characteristics (n=369)

		n	%
Age	18-25	55	14.9
	26-30	87	23.6
	31-35	84	22.8
	36-40	39	10.7
	40 and above	104	28.2
Gender	Men	157	42.5
	Women	212	57.5
Marital status	Single	181	49.1
	Married	188	50.9
Education	Undergraduate	60	16.3
	Graduate	142	38.5
	Master	56	15.2
	PhD	37	10.0
	Medical Doctor / Dentist	74	20.0
Profession	Medical Doctor	70	19.0
	Nurse	164	44.4
	Dentist	39	10.5
	Midwifery	20	5.4
	Physiotherapist	16	4.3
	Paramedic	36	9.8
	Health Manager	24	6.6
Organisation type	Public	208	56.4
	Private	112	30.3
	University	49	13.3
Service	Emergency	16	4.3
	Operating Room	10	2.7
	Administrative Department	122	33.0
	In-Patient Service	158	42.8
	Out-Patient Service	37	10.0
	Intensive Care Unit	26	7.0

Validity of Scale

Content validity and construct validity were tested in order to evaluate the validity of the scale.

Content Validity

The scope validity of the scale was evaluated in accordance with the Davis technique (Karagöz, 2017; Stevens and Mahwah, 2002). As a result of the content validity evaluation, it was seen that the CVR of all items varied between 0.84 and 1 and the CVI was 0.90. The scale was found to be valid in terms of scope with 20 items (Table 2).

Table 2. Content Validity

Items	Expert Opinion				CVR	Items	Expert Opinion				CVR
	a	b	c	d			a	b	c	d	
1	1		2	10	0.92	**11**		1	2	10	0.92
2	1	1	2	9	0.84	**12**	1		1	11	0.92
3	1		2	10	0.92	**13**	1		1	11	0.92
4	2		2	9	0.84	**14**	1		1	11	0.92
5	1		3	9	0.92	**15**	2			11	0.84
6		1	6	6	0.92	**16**	2		2	8	0.84
7	1	1	2	9	0.84	**17**	1	1	3	8	0.84
8			1	12	1	**18**	2			11	0.84
9			1	12	1	**19**	1		1	11	0.92
10				13	1	**20**			3	10	1
CVI: 0.90											

CVR-content validity ratio, CVI-content validity index, a-Not relevant, b-Much correction required, c-Relevant but little correction required, and d-Relevant

Construct Validity

The construct validity of the scale was evaluated with EFA, CFA and convergent-discriminant validity, and the suitability of the scale for factor analysis was determined with Bartlett's Sphericity test and KMO tests (Korku, 2002; Karakoç and Dönmez, 2014).

Global Nurs. Insights 2025,12(2), 100-108

Suitability for Factor Analysis

Bartlett Sphericity test and KMO tests were conducted to evaluate the applicability of factor analysis20,22. In order to perform factor analysis, KMO measurement results must be close to 1 (Karakoç and Dönmez, 2014). The KMO coefficient of the scale - sample adequacy measurement result was calculated as 0.87 and it was seen that it was suitable for factor analysis (Table.3).

Exploratory Factor Analysis

The factor loadings calculated to reveal the factor structure and loadings of the scale items were between 0.50 and 0.89, and the total variance explained in the scale was calculated as 55.48%, and it was seen that the scale consisted of a single dimension (Karakoç and Dönmez, 2014) (Table 3).

Confirmatory Factor Analysis

According to the EFA result, the validity of the scale, which was found to have strong values, was tested with confirmatory factor analysis. The fit indices of the scale were found as χ2 /df (4.68), RMSEA (0.09), NFI (0.90), NNFI (0.90), SRMR (0.084) and CFI (0.95) (Karagöz, 2017) (Table 3 and Figure 1).

Table 3. Construct Validity

KMO Coefficient – Sample Adequacy			0.87
Bartlett' s Test	**Chi-square**		7753.57
	Df		190
	P		<0.000
Items	**Factor Loadings**	**Self Value**	**Variance Explained (%)**
1	0.74	11.098	55.489
2	0.83		
3	0.74		
4	0.75		
5	0.70		
6	0.59		
7	0.83		
8	0.77		
9	0.50		
10	0.74		
11	0.84		
12	0.69		
13	0.71		
14	0.75		
15	0.91		
16	0.85		
17	0.67		
18	0.59		
19	0.57		
20	0.89		
	Good Fit	**Acceptible Fit**	**Model**
χ2 /df	1≤ χ2 /df ≤3	3<χ2 /df ≤ 5	4.68
RMSEA	0 ≤ RMSEA ≤ 0.05	0.05 <RMSEA ≤ 0.10	0.09
NFI	0.95 ≤ NFI ≤ 1	0.90 <NFI < 0.95	0.90
NNFI	0.95 ≤ NFI ≤ 1	0.90 <NNFI < 0.95	0.90
SRMR	0 ≤ SRMR < 0.05	0.05 ≤ SRMR < 0.10	0.08
CFI	0.97 ≤ CFI ≤ 1	0.95 ≤ CFI < 0.97	0.95

KMO-Keiser-Mayer-Olkin, χ2 /df : The ratio of the chi-square statistic to the respective degrees of freedom, RMSEA: Root mean square error of approximation, NFI: The normed fit index , NNFI: The (Non) normed fit index, SRMR: Standardized root mean square residual, CFI: Comparative fit indices.

Global Nurs. Insights 2025,12(2), 100-108

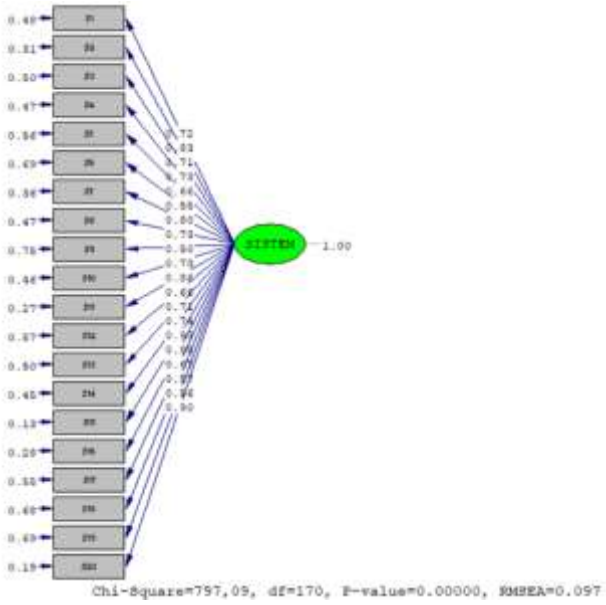


Figure 1: Scale Standardized Model

Convergent – Discriminant Validity

In order to investigate the validity of the measurement model, convergent and discriminant validity were examined (Erkorkmaz et al., 2013). In the analyses conducted to evaluate the convergent and discriminant validity of the scale, the AVE value of the scale was found to be 0.53 and the CR value was found to be 0.95 (Table 4)

Table 4. Convergent and Discriminant Validity

Items	FL	AVE	CR
1. I seek everyone's view of the situation.	0.72	0.53	0.95
2. I look beyond a specific event to determine the cause of the problem.	0.83		
3. I think understanding how the chain of events occurs is crucial.	0.71		
4. I include people in my work unit to find a solution.	0.73		
5. I think recurring patterns are more important than any one specific event.	0.66		
6. I think of the problem at hand as a series of connected issues.	0.55		
7. I consider the cause and effect that is occurring in a situation.	0.80		
8. I consider the relationships among coworkers in the work unit.	0.73		
9. I think that systems are constantly changing.	0.50		
10. I propose solutions that affect the work environment, not specific individuals.	0.73		
11. I keep in mind that proposed changes can affect the whole system.	0.86		
12. I think more than one or two people are needed to have success.	0.66		
13. I keep the mission and purpose of the organization in mind.	0.71		
14. I think small changes can produce important results.	0.74		
15. I consider how multiple changes affect each other.	0.93		
16. I think about how different employees might be affected by the improvement.	0.88		
17. I try strategies that do not rely on people's memory.	0.67		
18. I recognize system problems are influenced by past events.	0.57		
19. I consider the past history and culture of the work unit.	0.56		
20. I consider that the same action can have different effects over time, depending on the state of the system.	0.90		

FL-Factor Loadings, AVE-Explained Variance, CR-Composite Reliability

Reliability of the Scale

The reliability of the scale was tested with test-retest reliability and internal consistency analysis (Karakoç and Dönmez, 2014; Ylinen and Gullkvist, 2014).

Test-Retest Analysis

In order to evaluate the stability and reliability of the scale over time, the scale draft was applied to a group of 30 participants with similar characteristics to the sample group twice with a 4-week interval (Karakoç and Dönmez, 2014; Ylinen and Gullkvist,

2014). The data were evaluated with paired group t-test and Pearson correlation. The intraclass correlation coefficient (ICC) determined in the scale was found to be 0.99 ($p<0.01$). It was seen that the same or similar results were obtained in both measurements (Table.5).

Internal Consistency Analysis

Cronbach alpha internal consistency analysis was performed to evaluate the internal consistency of the scale. The Cronbach alpha value of the scale was found to be 0.95 (Table 5).

Table 5. Reliability

	Test		**Re-Test**	
Total Scale	**Mean±SD**	**Min-Max (Median)**	**Mean±SD**	**Min-Max (Median)**
	4.03±0.74	1-5 (4.03)	4.05±0.75	1-5 (4.2)
	Corelation		**ICC**	
Total Scale	**r**	**p**	**r**	**p**
	0.99	0.001*	0.99	0.001*
Total Scale Cronbach's Alpha			**0.95**	

SD: Standart Deviation, Min: Minimum, Max: Maximum, ICC: Intraclass correlation coefficient

DISCUSSION

Systems thinking is a critical component of quality improvement learning and has an important place in health systems. The need for systems thinking, which is considered as the main element in the joint and balanced functioning of health systems and forms the basis for the system's competencies, is increasingly emphasized. It is thought that the systems thinking scale consisting of 20 items will contribute to the implementation of the systems approach in clinical practice based on systems thinking and even in the education of health professionals. Information needs to be measured and data obtained for the production and development of scientific knowledge. Scales are used for this purpose and they are measurement and data collection tools obtained within the framework of certain rules for the classification, ranking and quantification of the desired characteristics. One of the most important points in the use of scales is to be able to obtain accurate data and for this reason, scales are expected to be scientifically valid and reliable (Korku, 2002). This research was conducted to test the validity and reliability of the systems thinking scale in Turkish and the findings were discussed in the light of the literature. Validity refers to the suitability of the scale for the subject and feature it is intended to measure and the generalizability of the results obtained (Bolarinwa, 2015; Ylinen and Gullkvist, 2014). Davis technique was used in the assessment of the scope validity of the scale.

Construct validity was assessed with exploratory and confirmatory factor analyses and convergent-discriminant validity, and the suitability of the data for factor analysis was tested with Bartlett Sphericity test and Keiser-Mayer-Olkin tests (Bolarinwa, 2015; Stevens and Mahwah, 2002). According to the Davis technique, the content validity indexes of the scale items and the scale content validity ratio should be higher than 0.80 (Bolarinwa, 2015; Yeşilyurt and Çapraz, 2018). The content validity ratios of the scale items are between 0.84 and 1 and the content validity ratio is 0.90, and it was observed that all scale items and the total scale provided content validity. Bartlett Sphericity test and Keiser-Mayer-Olkin tests evaluate the suitability of the scale for factor analysis. The KMO coefficient should be 0.60 and above and the Bartlett significance test should be statistically significant at ($p<0.05$) level (Davis, 1992; Korku, 2002). The KMO coefficient was 0.87 and the scale was found to be suitable. EFA evaluates the structure of a scale and determines the different factors it will measure, and the variance ratios of the scale and sub-factors are calculated. The size of the ratios obtained reveals the strength of the factor structure and the results should be 40% to 60%26. For an item to measure the structure or factor well, this value is expected to be 0.40 or above (Korku, 2002; Yurdabakan and Çüm, 2017). In the study, the total variance explained in the scale was 55.48% and the factor loadings of all items ranged between 0.50 and 0.89 and were found to be suitable above 0.40.

CFA aims to confirm the obtained scale structure and factors, and scale fit indices are calculated with the scale confirmatory factor analysis model. Fit index is called $\chi 2$, $\chi 2/df$, GFI, CFI, RMSEA, SRMR and is evaluated as "acceptable" and "normal" fit (Karagöz, 2017; Bolarinwa, 2015). In the study, the fit indices of the scale were found to be $\chi 2/df$ (4.68), RMSEA (0.097), NFI (0.90), NNFI (0.90), SRMR (0.08) and CFI (0.95) at acceptable fit values. The confirmatory factor analysis of the scale is statistically significant and valid. Convergent-discriminant validity shows the tendency of all items of the scale to confirm each other. Ylinen and Gullkvist (2014) stated that convergent validity can be assessed with CR and AVE (Karagöz, 2017). CR indicates structural consistency, and AVE indicates the relative variance attributed to the structure. In each structure evaluated, CR should be 0.6 and above and AVE should be 0.5 and above (Bogozzi and Yi, 1998; Gökdemir and Yılmaz, 2023). CR was found to be 0.95 and AVE was found to be 0.53, above the expected values. Reliability shows how consistently the scale measures the variable it aims to measure or how error-free the measurement results are. This refers to the stability of the values obtained from measurements made with the measurement tool under the same conditions (Azwa et al., 2016; Karakoç and Dönmez, 2014). In the study, the reliability of the scale was tested with test-retest and internal consistency analysis. In the test-retest method, the scale is applied to the same sample group more than once, under the same conditions and at the same times. The obtained results are tested with correlation analysis and the reliability coefficient is calculated (Gökdemir and Yılmaz, 2023). In case the measured feature is a continuous

variable, and the developed scale is an equal interval or ratio scale, Pearson correlation analysis is used, and the correlation coefficient is calculated. The correlation coefficient examines the degree and direction of the relationship between two variables and varies between -1 and +1. A result of +1 indicates a positive and perfect relationship (Azwa et al., 2016; Bolarinwa, 2015). The ICC values determined in the scale were found to be close to 1 with r=0.99, p<0.001. It was observed that the same or similar results were obtained in both measurements. Internal consistency evaluates the homogeneity of the items forming the structure to be measured among themselves and the level to which they measure the concept to be measured. Cronbach Alpha coefficient is frequently used for this purpose. For a scale to be reliable, the Cronbach Alpha value must be between 0.70 and 0.99 (Bolarinwa, 2015; Tavakol and Dennick, 2011). Cronbach's Alpha coefficient in the study was found to be 0.95, which is consistent with the literature.

CONCLUSIONS

The system thinking scale adapted to Turkish consists of 20 items and a single dimension. The scale, which has good psychometric properties in the adapted language, can be used to evaluate system thinking in health systems. It is recommended that the scale be tested in larger samples and with different variables. It is thought that the use of the scale will create awareness for ensuring system security in health systems and contribute to the orientation of considering the system as a whole.

Ethics Committee Approval: İstanbul Kent University Ethical Committe, 18.01.2024, 2024/2
Conflict of Interest: Author declares no conflict of interest
Funding: Author declares no financial support
Exhibitor Consent: It was taken on the first part of online data collection form.
Author contributions
Study design: GA
Data collection: GA
Literature search: GA
Drafting manuscript: GA
Acknowledgment: I would like to thank to all health professionals participated to the study.

REFERENCES

Arnold, R.D., & Wade, J.P. (2015). A Definition of Systems Thinking: A Systems Approach, Procedia Computer Science, 44, 669 – 678. https://doi.org/10.1016/j.procs.2015.03.050.

Azwa Ambad, S.N., & Abdul Wahab, K. (2016). The relationship between corporate entrepreneurship and firm performance: evidence from malaysian large companies. International Journal of Business and Society, 17, 259– 80.

Bagozzi, R., & Yi, Y. (1998). On the Evaluation of Structural Equation Models. Journal of the Academy of Marketing Sciences, 16:74- 94. http://dx.doi.org/10.1007/BF02723327.

Baştürk, S., Dönmez, G., Dicle, A. & Baştürk, S. (2013). Validity and Reliability. Scientific Research Methods Book. Vize Publishing.

Batalden, P. B., & Mohr, J. J. (1997). Building knowledge of health care as a system. Quality management in health care, 5(3), 1–12.

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. The Nigerian postgraduate medical journal, 22(4), 195–201. https://doi.org/10.4103/1117-1936.173959

Bozkurt, A., & Yılmaz, M. (2024). Adaptation of cancer screening perception scale to Turkish: validity and reliability study. HUJNF, 11(2), 135-142. https://doi.org/10.31125/hunhemsire.1344212.

Burmakoğlu, S., Kıdak, L.B., Sur, H., & Demir, H. (2016). System approach and system dynamics applications in health field: a biometric analysis. Hacettepe Health Administration Journal, 19(4), 465- 479.

Çapık, C., Gözüm, S., & Aksayan, S. (2018). Cross-cultural scale adaptation stages, language and culture adaptation: updated guide. FNJN, 26(3), 199- 210. https://doi.org/10.26650/FNJN397481.

Davis, L. (1992). Instrument review: Getting the most from a panel of experts. Appl Nurs Res, 5, 194-7. https://doi.org/10.1016/S0897-1897(05)80008-4.

Dolansky, M. A., Moore, S. M., Palmieri, P. A., & Singh, M. K. (2020). Development and Validation of the Systems Thinking Scale. Journal of general internal medicine, 35(8), 2314–2320. https://doi.org/10.1007/s11606-020-05830-1.

Erkorkmaz, Ü., Etikan, İ., Demir, O., Özdamar, K., & Sanisoğlu, S.Y. (2013). Confirmatory factor analysis and fit indices. Turkiye Clinics J Med Sci, 33(1), 210- 23. https://doi.org/10.5336/medsci.2011-26747.,

Gökdemir, F., & Yılmaz, T. (2023). Processes of using , modifying, adapting and developing likert t ype scales. J Nursology. 26(2), 148- 160. https://doi.org/10.5152/JANHS.2023.22260.

Hair, J., Anderson, R.E., Tatham, R.L., & Black, W.C. (1995). Multivariate data analysis (4th ed.) New Jersey: Prentice-Hall Inc.

Karagöz, Y. (2017). SPSS and AMOS Applied Quantitative-Qualitative-Mixed Scientific Research Methods and Publication Ethics. Nobel Publishing Distribution

Karakoç, F.Y., & Dönmez, L. (2014). Basic principles in scale development studies. Medical Education World.

Korku, C. (2002). Psychometric properties of the Turkish version of the work-related quality of life scale for nurses. Hacettepe Health Administration Journal, 25(2), 287- 300.

World Health Organization. (2020). Patient safety incident reporting and learning systems: Technical report and guidance. World Health Organization. https://iris.who.int/handle/10665/334323

Phillips, J. M., Stalter, A. M., Winegardner, S., Wiggs, C., & Jauch, A. (2018). Systems thinking and incivility in nursing practice: An integrative review. Nursing forum, 10.1111/nuf.12250. Advance online publication. https://doi.org/10.1111/nuf.12250.

Plack, M. M., Goldman, E. F., Scott, A. R., Pintz, C., Herrmann, D., Kline, K., Thompson, T., & Brundage, S. B. (2018). Systems Thinking and Systems-Based Practice Across the Health Professions: An Inquiry Into Definitions, Teaching Practices, and Assessment. Teaching and learning in medicine, 30(3), 242–254. https://doi.org/10.1080/10401334.2017.1398654.

Rusoja, E., Haynie, D., Sievers, J., Mustafee, N., Nelson, F., Reynolds, M., Sarriot, E., Swanson, R. C., & Williams, B. (2018). Thinking about complexity in health: A systematic review of the key systems thinking and complexity ideas in health. Journal of evaluation in clinical practice, 24(3), 600–606. https://doi.org/10.1111/jep.12856.

Stevens, J.P. & Mahwah, N.J. (2002). Applied multivariate statistics for the social sciences. Lawrence Erlbaum.

Tabachnick, B.G., & Fideli, L.S. (2018). Using Multivariate Statistics (7.Edition). Ally And Bacon.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha, International Journal of Medical Education, 2, 53– 5. https://doi.org/10.5116/ijme.4dfb.8dfd.

Tetuan, T., Ohm, R., Kinzie, L., McMaster, S., Moffitt, B., & Mosier, M. (2018). Does systems thinking ımprove the perception of safety culture and patient safety? Journal of Nursing Regulation, 8(2), 31– 39. https://doi.org/10.1016/S2155-8256(17)30096-0.

Yeşilyurt, S., & Çapraz C. (2018). A roadmap for content validity used in scale development studies. Erzincan University Faculty of Education Journal, 20(1), 251- 264. https://doi.org/10.17556/erziefd.297741.

Ylinen, M., & Gullkvist, B. (2014). The effects of organic and mechanistic control in exploratory and exploitative innovations, Management Accounting Research, 25(1), 93-112. https://doi.org/10.1016/j.mar.2013.05.001.

Yurdabakan, İ., & Çüm S. (2017). Scale Development in Behavioral Sciences, TJFMPC, 11(2), 108- 126. https://doi.org/10.21763/tjfmpc.317880.

Zivadinovic, K.N. (2004). Utvrđivanje osnovnih karakteristika proizvoda primjenom faktorske analize [Defining the basic product attributes using the factor analysis]. Ekonomski pregled, 55, 952– 966.