

The development and validation of the Midwives' Empathetic Tendency and Skill Intention Scale: A methodological study

Tuba BUYUKKAL^a, Fatma AY^b

ABSTRACT

High levels of empathetic skills and tendencies among midwives are associated with improved quality of care. The Theory of Planned Behavior (TPB) provides a useful framework for examining midwives' intentions regarding empathetic tendencies and skills, offering a theoretical basis for understanding and strengthening empathetic care in midwifery practice. The present study was undertaken to develop and validate a instrument for measuring midwives' intentions regarding empathetic tendencies and empathic skills based on the TPB. In the first phase of the study, a scale development process was conducted based on the TPB. A comprehensive literature review was undertaken, resulting in an initial item pool of 44 items. Following expert review, a 26-item draft scale was generated. The construct validity of the instrument was examined through both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). The study was completed with a total sample of 287 midwives. EFA yielded a three-factor structure comprising 15 items. Factor 1 (Attitude Toward the Behavior) explained 45.937% of the total variance, Factor 2 (Subjective Norm) explained 13.339%, and Factor 3 (Perceived Behavioral Control) explained 8.851%, with the three-factor model accounting for 68.127% of the total variance. The scale demonstrated high internal consistency, with an overall Cronbach's alpha coefficient of 0.892. CFA results indicated acceptable model fit indices, supporting the validity of the three-factor structure. The Midwives' Empathetic Tendency and Skill Intention Scale is a valid and reliable instrument that can be used as an effective measurement tool in midwifery research and practice.

Keywords: Empathy, validity, reliability, scale

Ebelerin Empatik Eğilimi ve Beceri Niyeti Ölçeği'nin geliştirilmesi ve geçerliliğinin doğrulanması: Metodolojik bir çalışma

Ebelerin empatik becerileri ve eğilimlerinin yüksek düzeyde olması bakım kalitesinin artmasıyla ilişkilidir. Planlanmış Davranış Teorisi (PDT), ebelerin empatik eğilimleri ve becerilerine ilişkin niyetlerini incelemek için yararlı bir çerçeve sunarak, ebelik uygulamalarında empatik bakımı anlamak ve güçlendirmek için teorik bir temel sağlamaktadır. Bu çalışma, PDT bağlamında ebelerin empatik becerisi ve eğilim niyetlerini değerlendirmek üzere geçerli ve güvenilir bir ölçüm aracı geliştirmek amacıyla yapılmıştır. Çalışmanın ilk aşamasında, PDT temelinde bir ölçek geliştirme süreci yürütülmüştür. Kapsamlı bir literatür taraması yapılmış ve sonuçta 44 maddeden oluşan bir ilk madde havuzu belirlenmiştir. Uzman incelemesinin ardından, 26 maddelik bir taslak ölçek oluşturulmuştur. Ölçeğin yapı geçerliliğini değerlendirmek için Açıklayıcı Faktör Analizi (AFA) ve Doğrulayıcı Faktör Analizi (DFA) yapılmıştır. Çalışma, toplamda 287 ebe örnekleme ile tamamlanmıştır. AFA ile 15 madde ve üç faktörden oluşan bir yapı ortaya çıkmıştır. Faktör 1 (Davranışa Yönelik Tutum) toplam varyansın %45,937'sini, Faktör 2 (Öznel Norm) %13,339'unu ve Faktör 3 (Algılanan Davranış Kontrolü) %8,851'ini açıklarken, üç faktörlü model toplam varyansın %68,127'sini açıklamaktadır. Ölçeğin Cronbach alfa katsayısı 0,892 olup yüksek iç tutarlılık sergilemektedir. DFA sonuçları bakıldığında kabul edilebilir model uyum indeksleri göstermekte ve bu da üç faktörlü yapının geçerliliğini desteklemektedir. "Ebelerin Empatik Eğilim ve Beceri Niyeti Ölçeği", ebelik araştırmaları ve uygulamalarında kullanılabilecek, geçerli ve güvenilir bir ölçüm aracıdır.

Anahtar Kelimeler: Empati, geçerlilik, güvenilirlik, ölçek

Geliş Tarihi: 02.04.2026

Kabul Tarihi: 31.07.2026

^aIstinye University, Faculty of Health Sciences, Department of Midwifery, İstanbul, Türkiye, e-posta: buyukkaltuba@gmail.com / tuba.buyukkal@istinye.edu.tr ORCID: 0000-0002-7156-4706

^bIstanbul University-Cerrahpaşa, Faculty of Health Sciences, Department of Midwifery, İstanbul, Türkiye, e-posta: fatma.ay@iuc.edu.tr ORCID: 0000-0001-9491-0731

Sorumlu Yazar/Correspondence: Fatma Ay e-posta: fatma.ay@iuc.edu.tr

*Çalışma doktora tezinden türetilmiştir.

Atıf/Citation: Buyukkal T, Ay F. The development and validation of the Midwives' Empathetic Tendency and Skill Intention Scale: A methodological study. Journal of Health and Life Science 2026;8(2):89-98.

INTRODUCTION

Empathy is one of the most studied concepts in the field of psychology throughout history; it is a multidimensional structure that expresses an individual's ability to focus on the person in front of them at an emotional, cognitive, and behavioral level, to understand them, and to reflect this understanding appropriately.¹ This concept, which originates from the Greek term "empathia," meaning "to feel inside," was first defined in the literature in 1897 by Lipps as "Einfühlung," meaning "feeling inside another." It was later explained as "Mitgefühl," meaning "feeling with another," and thus acquired its current meaning. Carl Rogers, one of the pioneers of humanistic psychology, defined empathy as the process of understanding another person's inner world from their perspective and accepting this understanding without judgment.^{1,2}

Empathy stands out as an important professional competency for midwifery, one of the professions that is closely intertwined with people, provides direct services to people, and requires effective communication skills. Empathy is considered a fundamental component in midwifery undergraduate education and professional development processes.³⁻⁵ Developing midwives' empathic skills has been shown to have positive effects on their ability to provide individual-centered and personalized care. Midwives who work with an empathic approach have been found to establish stronger communication with individuals, create an appropriate therapeutic environment, and improve the quality of care. Therefore, continuously supporting and developing empathic skills is of great importance for the sustainability of qualified midwifery services.^{5,6}

The Theory of Planned Behavior (TPB) was introduced into the literature after the "Theory of Reasoned Action" developed by Ajzen and Fishbein in 1975. Aiming to explain, predict, and identify the factors influencing an individual's behavior, the TPB argues that the person's intention plays a role in performing that behavior. Within the framework of the TPB, behavioral intention is explained by three interrelated determinants: subjective norms, attitudes toward the behavior, and perceived behavioral control.⁷⁻⁹ A review of the literature reveals that scales exist to determine the level of empathy skills and tendencies that are valid and reliable across different professional groups.¹⁰ However, a validated scale to assess the extent to which midwives' empathic tendencies and empathic skill intentions influence their behavior has not yet been developed. Accordingly, the present study seeks to design a valid and reliable instrument for evaluating midwives' empathic tendencies and skill intentions within the framework of the TPB.

Materials And Methods

Design

This research was conducted using methodological design. The data for the research were gathered through an online survey conducted between March and December 2023.

Data Collection

The six-stage process outlined by De Vellis (2017) was followed in the scale development process.¹¹ Taking these stages into account, the scale development process was carried out as follows.

1. Literature Review: In the first stage, the structures to be measured and related theoretical frameworks were determined; concepts of empathy, tendency, and intention were addressed within the context of the TPB, and a comprehensive review of the existing literature was undertaken. This review indicated that there is no original scale specifically designed to measure midwives' empathetic skills and tendency levels within the framework of the TPB. Therefore, a new scale development process was initiated in the first phase of the study.¹²

2. Item Development: In order to create an item pool to measure the identified constructs, individual interviews were conducted with 10 midwives independent of the research universe. During the interviews, nine open-ended questions, created using the literature and including TPB dimensions, were asked. Based on the responses obtained, frequently repeated statements were analyzed to create a 44-item pool related to empathic skills and intention to empathize.¹¹

3. Expert Opinion: Expert review was carried out to assess the content validity, relevance, and linguistic clarity of the initial item pool. A panel of seven experts participated in this process, including four specialists in midwifery, one in obstetrics and gynecology, one in medicine and one in psychology. Item content validity was evaluated using the Lawshe method by calculating the Content Validity Ratio (CVR). Based on the evaluations of the seven experts, 18 items with CVR values below the Lawshe critical threshold (CVR = 1.00) were eliminated, resulting in a preliminary 26-item scale.¹³⁻¹⁵

4. Pilot Application: A pilot application was conducted with 30 randomly selected midwives to assess whether the scale items were clear and understandable. After the pilot application, no need for corrections in terms of language and comprehensibility was identified.

5. Application: The finalized 26-item draft scale was then administered to a sample of 131 midwives, corresponding to at least five times the number of items. This approach is based on the recommendation that the sample size in scale development studies should be between 5 and 10 times the total number of items.^{16,17} In addition to the scale, participants' sociodemographic information was obtained using a researcher-designed Personal Information Form.

Following the finalization of the scale, test-retest reliability was evaluated by administering the finalized instrument to 50 midwives twice, with a two-week interval between administrations. The findings indicated satisfactory temporal stability of the scale.

6. Analysis: All statistical analyses were carried out using SPSS version 23. Following data analysis, the final version of the scale was established. Exploratory Factor Analysis (EFA) was conducted to assess construct validity. In this process, principal component analysis was applied for factor extraction, while direct oblimin rotation was preferred for factor rotation. Internal consistency reliability was assessed using Cronbach's alpha. Descriptive analyses included frequencies and percentages for categorical variables, while continuous variables were expressed as mean $\pm$ standard deviation or as median (range), according to the distribution of the data. Statistical significance was accepted at $p < 0.05$.

Participants

During the EFA stage, data were obtained from 131 midwives, while 106 midwives participated in the Confirmatory Factor Analysis (CFA) stage. For the test-retest procedure, a separate group of 50 midwives was included, resulting in a total sample of 287 participants. Additionally, a pilot study was carried out with 30 midwives to assess the clarity and comprehensibility of the scale items; these participants were not included in the main sample.

The sample group included midwives working in Turkey who held a midwifery license approved by the Ministry of Health. Midwives who did not work as clinical midwives or who had different roles such as education or management and who were not Turkish citizens were excluded from the study.

Ethical approval was obtained from the Istanbul University-Cerrahpaşa Non-Interventional Clinical Research Ethics Committee. The study was conducted in accordance with the principles of the Declaration of Helsinki. Furthermore, upon accessing the online survey, participants first encountered an informed consent form detailing the purpose of the study, who

conducted it, and data security measures. Only after providing consent could users view the survey and scale questions.

RESULTS

Descriptive Characteristics of Participant Samples

Looking at the participants' descriptive characteristics, the average age was 32.39 ± 8.25 , and 67.9% were college graduates. 31.3% of participants had two siblings, and 79.4% grew up in a nuclear family. The percentage of married participants was 56.5%, while 46.6% had children. Of those with children, 25.2% had two children. Only 32.1% worked in teaching hospitals. While 23.7% of participants worked in primary health care services, the remaining large proportion worked in different units. 57.3% of participants stated that they chose the midwifery profession willingly, and 78.6% stated that they enjoy/willingly perform their profession. When working hours were evaluated, it was determined that 48.9% worked 41-60 hours per week. Sixty-seven-point nine percent of participants stated that they had primarily not cared for a relative in the last six months (Table 1).

Construct Validity (Exploratory Factor Analysis)

EFA was conducted to determine the construct validity and underlying factor structure of the scale. As correlations among factors were anticipated, a direct oblimin rotation method was employed.^{18,19}

The Kaiser-Meyer-Olkin (KMO) coefficient was used to evaluate sampling adequacy, with values above 0.50 considered acceptable.^{19,20} The overall KMO value for the scale was calculated as 0.878, while the item-level KMO value was found to be 0.913 (Table 2), indicating that the sample size was adequate for factor analysis.^{19,20}

In factor analysis, items loading on more than one factor are considered problematic; specifically, if the difference between factor loadings is less than 0.10, the item should be removed.^{19,20} Accordingly, items 1, 2, and 10 were excluded during the first analysis, item 6 during the second analysis, and item 14 during the third analysis due to cross-loading on multiple factors. Since a minimum of three items must be present in a sub-dimension¹⁸⁻²⁰, item 3 and item 4 were removed from the scale as they were found in a single sub-dimension in the third analysis, and item 9 and item 18 were removed as they were found in a single sub-dimension in the fifth analysis.

Table 1. Demographic and occupational characteristics (n=131)

Age	Mean $\pm$ SD	32.39 $\pm$ 8.25		
	Median (Min-Max)	31 (22 - 63)		
			Frequency	Percentage
Educational status	Bachelor's degree		89	67.9
	Master's degree		25	19.1
	Doctorate		17	13
Marital status	Single		57	43.5
	Married		74	56.5
Children	Yes		61	46.6
	No		70	53.4
Number of children	None		70	53.4
	1		25	19.1
	2		33	25.2
	3		3	2.3
Type of family raised in	Nuclear family		104	79.4
	Extended family		23	17.6
	Broken family		4	3.1
Number of siblings	None		5	3.8
	1		37	28.2
	2		41	31.3
	3		27	20.6
	4 and above		21	16
Voluntarily choosing the midwifery profession	Yes		75	57.3
	No		46	35.1
	Undecided		10	7.6
Doing the job with enthusiasm/willingness	Yes		103	78.6
	No		8	6.1
	Undecided		20	15.3
Institution	Primary health care services		31	23.7
	Teaching and research hospital		42	32.1
	Public hospital		36	27.5
	Private hospital		22	16.8
Department	Delivery room		26	19.8
	Obstetrics and gynecology department		32	24.4
	Primary health care services		31	23.7
	Newborn / Infant		19	14.5
	Other units (outside the midwifery field and training unit)		23	17.6
Weekly working hours	30-40 hours		56	42.7
	41-60 hours		64	48.9
	More than 61 hours		9	6.9
	Less than 29 hours		2	1.5
Having cared for a relative in the last six months	Yes		42	32.1
	No		89	67.9

In the fourth analysis, the correlation matrix table of the scale, which was grouped under four factors, was examined. Following this evaluation, items 19 and 20 were excluded from the scale due to their lack of meaningful association with the other items.¹⁸⁻²⁰ After removing items 1, 2, 3, 4, 6, 9, 10, 14, 18, 19, and 20

from the model, the KMO value remained at 0.878, and the Bartlett's test of sphericity yielded a chi-square value of $\chi^2 (105) = 1213.385$ ($p < 0.001$). These findings confirm that the dataset is appropriate for factor analysis (Table 2).

Table 2. KMO and Bartlett Sphericity Test results

Kaiser-Meyer-Olkin Sample Adequacy Measure	0.878
Bartlett's Sphericity Test	Chi square
	1213.385
	Degree of freedom
	105
	Significance
	0.000

It was also observed that all extraction values of the 15-item scale exceeded 0.30 and that the diagonal values in the anti-image correlation matrix were above 0.50.¹⁸⁻²⁰ The analysis revealed a three-factor structure. Factor

1 (Attitude Toward the Behavior) accounted for 45.937% of the total variance, Factor 2 (Subjective Norm) for 13.339%, and Factor 3 (Perceived Behavioral Control) for 8.851%. Altogether, these

three factors explained 68.127% of the total variance (Table 3). Factor 1 comprised Items 11, 12, 13, 15, 16, and 17; Factor 2 included Items 5, 7, and 8; and Factor 3 consisted of Items 21, 22, 23, 24, 25, and 26. Examination of the conceptual content of the items loading on each factor indicated that Factor 1 represented perceived behavioral control, Factor 2 reflected subjective norm, and Factor 3 corresponded to attitude toward the behavior.

According to the scale's analysis results, Eigen (eigenvalue) and initial common values were examined. The lowest initial common value among the items was found to be 0.444 and the highest was 0.843. The values were found to be appropriate as they were above 0.30. At the same time, the mean, variance, correlation, and Cronbach's alpha values for each item were calculated. In terms of reliability, the overall Cronbach's alpha coefficient of the scale was calculated as 0.892, indicating a high level of internal consistency.

Table 3. Exploratory factor analysis and internal consistency results for the scale

	Factor 1	Factor 2	Factor 3	Extraction
I-11	0.862			0.736
I-12	0.757			0.738
I-13	0.889			0.716
I-15	0.735			0.725
I-16	0.755			0.586
I-17	0.819			0.701
I-5		0.751		0.565
I-7		0.826		0.766
I-8		0.923		0.843
I-21			0.805	0.618
I-22			0.786	0.727
I-23			0.733	0.803
I-24			0.657	0.664
I-25			0.676	0.444
I-26			0.699	0.586
Eigen	6.891	2.001	1.328	
VER	45.937	13.339	8.851	
CVER	45.937	59.276	68.127	
CA	0.907	0.810	0.855	
CA	0.892			

Table 4. Cronbach's α coefficient of scale

Factor analysis Cronbach's α Coefficient of Scale							
	Factor Loadings	Mean	SD	If the item is deleted, the average of the scale	If the item is deleted, the variance of the scale	Corrected item-total score correlation	If the Item Is deleted, the Cronbach α coefficient of the scale
Factor 1: Perceived behavioral control							
α : 0.907							
Mean: 4.43 $\pm$ 0.53							
% Var: 69.42							
Item 11	0.862	4.47	.586	59.91	46.530	.682	.882
Item 12	0.757	4.51	.625	59.87	45.945	.708	.881
Item 13	0.889	4.45	.610	59.93	46.788	.619	.884
Item 15	0.735	4.48	.637	59.90	45.613	.734	.880
Item 16	0.755	4.24	.795	60.14	45.319	.596	.884
Item 17	0.819	4.46	.636	59.92	46.148	.669	.882
Factor 2: Subjective norm							
α : 0.810							
Mean: 3.89 $\pm$ 0.68							
% Var: 72.71							
Item 5	0.751	3.76	.902	60.62	47.715	.305	.891
Item 7	0.826	4.05	.955	60.33	45.284	.477	.890
Item 8	0.923	3.86	.901	60.52	47.067	.360	.890
Factor 3: Attitude toward behavior							
α : 0.855							
Mean: 4.34 $\pm$ 0.59							
% Var: 61.70							
Item 21	0.805	4.18	.967	60.20	45.330	.466	.891
Item 22	0.786	4.39	.697	59.99	45.408	.685	.881
Item 23	0.733	4.43	.724	59.95	44.275	.782	.877
Item 24	0.657	4.49	.600	59.89	46.219	.705	.881
Item 25	0.676	4.16	.910	60.22	45.835	.460	.891
Item 26	0.699	4.44	.735	59.95	45.574	.627	.883

Cronbach's alpha values were 0.907 for Factor 1, 0.810 for Factor 2, and 0.855 for Factor 3. Item discrimination was assessed using the lower 27%–upper 27% group comparison, and a statistically significant difference was found between the total

scores of these groups ($p < 0.05$).¹⁸⁻²⁰ These results suggest that the scale demonstrates strong discriminative validity and high reliability (Table 4). The variance values of the scale with three subscales were examined, and the cumulative variance values of

the subscales were calculated. It was determined that the three subscales accounted for 68.127% of the entire scale. Since this value was expected to be above 50%, it was considered sufficient.

The convergent validity of the three-factor structure was assessed using the Average Variance Extracted (AVE) and Composite Reliability (CR) values. For Factor 1, the AVE and CR values were 0.638 and 0.9132, respectively. For Factor 2, the AVE and CR values were 0.6938 and 0.8709, respectively. For Factor 3, the AVE and CR values were 0.5215 and 0.8667, respectively. All factors achieved AVE values above the recommended threshold of 0.50 and CR values exceeding 0.70, demonstrating that the three-factor model has sufficient convergent validity and satisfactory composite reliability.²²

Finally, the correlations between the subscales found in the construct validity were examined. Low-to-moderate correlations were found between Factor 1

and Factor 2 and between Factor 2 and Factor 3, while a high correlation was found between Factor 1 and Factor 3. Since the correlation values were not above 0.80, they were deemed appropriate (Table 5).

Table 5. Relationship between the subdimensions of the Midwives' Empathetic Tendency and Skill Intention Scale

		Factor 1	Factor 2	Factor 3
Factor 1	<i>r</i>	1.000	.275	.657
	<i>p</i>		.001**	.000**
Factor 2	<i>r</i>		1.000	.254
	<i>p</i>			.003**
Factor 3	<i>r</i>			1.000
	<i>p</i>			

r= Spearman's Correlation

***p*<0,01

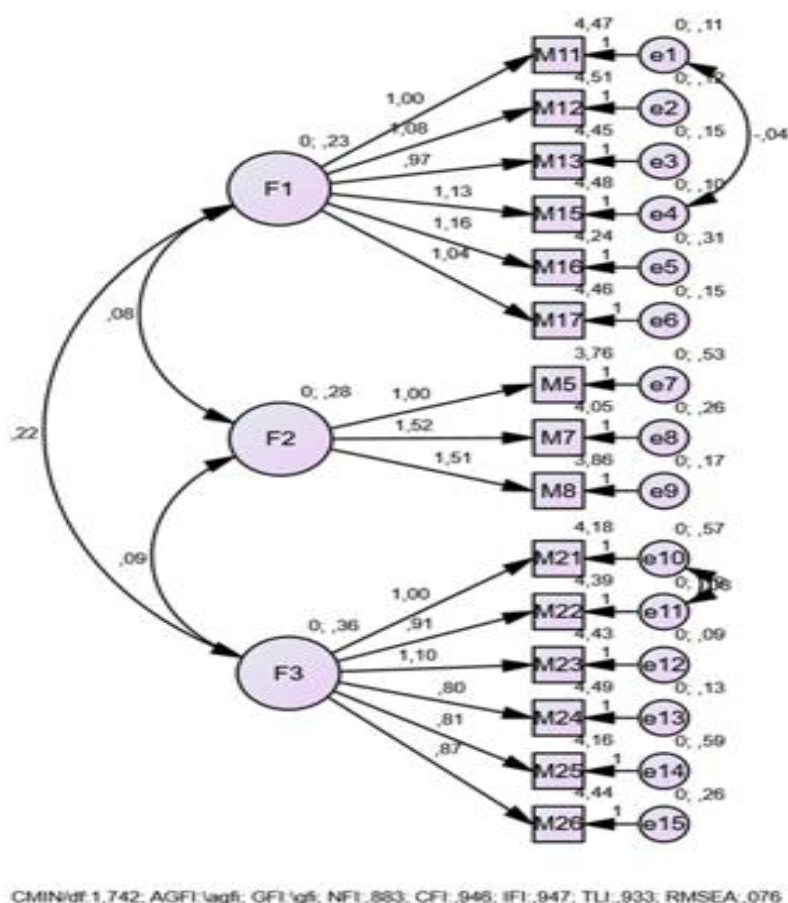


Figure 1. Confirmatory factor analysis. Standardized scores of the three-factor (F1, F2, F3) structure of the scale

Confirmatory Factor Analysis Stage

The 15-item factor structure identified through EFA was subsequently examined using CFA. Each item was

placed under the factors that emerged because of the analysis, and the necessary covariances were established. As a result of CFA, it was determined that the factor loadings of each item of the scale were

between 0.53 and 0.91. The text regarding the analysis results and improvements was reviewed. The literature suggests adding error covariance if items are under the same factor, have similar content, or contain similar statements²¹⁻²³ Based on the suggestions, covariance lines showing the relationships between e1 and e4 and between e10 and e11 were added. As a result, while there was no change in the factor loadings of the scale items, the correlations between the subscales were determined to be between 0.30 and 0.76. When reviewing the literature on CFA, it is stated that the factor loadings of each item should be at least 0.30 or 0.40, and the correlations between the sub-dimensions should be in the range of 0.30 to 0.70.²¹⁻²⁴ In this regard, the factor loadings of the scale items and the correlations between the subscales were found to be adequate (Figure 1).

Another value that should be examined based on the CFA results of the scale is the model fit indices. The literature contains many fit indices related to CFA and their fit scale ranges. Among these, the most preferred are CMIN/Sd, AGFI, GFI, CFI, NFI, IFI, TLI, and RMSEA values.^{21-23,25} The scale's goodness-of-fit values are detailed in Table 6. The values are acceptable and fall within good ranges.

To examine the temporal stability of the instrument, test-retest reliability was evaluated by administering the finalized scale to the same group of 50 participants on two separate occasions with a two-week interval. The association between the total scores from the first and second administrations was examined using Pearson's correlation coefficient. The findings revealed a strong, positive, and statistically significant correlation, supporting the stability of the scale over time ($r = 0.85$, $p < 0.001$).

Tablo 6. Fit Index Values obtained from confirmatory factor analysis

Measurement	Good Fit	Acceptable Fit	Model Fit Index Values
RMSEA	$0 < \text{RMSEA} < 0.05$	$0.05 \leq \text{RMSEA} \leq 0.10$	0.076 Acceptable fit
NFI	$0.90 \leq \text{NFI} \leq 1$	$0.85 \leq \text{NFI} \leq 0.89$	0.88 Acceptable fit
TLI	$0.95 \leq \text{NNFI} \leq 1$	$0.90 \leq \text{NNFI} \leq 0.95$	0.93 Acceptable fit
CFI	$0.97 \leq \text{CFI} \leq 1$	$0.95 \leq \text{CFI} \leq 0.97$	0.94 Acceptable fit
IFI	$0.97 \leq \text{IFI} \leq 1$	$0.95 \leq \text{IFI} \leq 0.97$	0.94 Acceptable fit
SRMR	$0 \leq \text{SRMR} \leq 0.05$	$0.05 \leq \text{SRMR} \leq 0.10$	0.051 Acceptable fit
RFI	$0.90 \leq \text{RFI} \leq 1$	$0.85 \leq \text{RFI} \leq 0.90$	0.85 Acceptable fit

RMSEA, Root Mean Square Error of Approximation; NFI, Normed Fit Index; TLI, Tucker-Lewis Index; CFI, Comparative Fit Index; IFI, Incremental Fit Index; SRMR, Standardized Root Mean Square Residual; RFI, Relative fit indices

DISCUSSION

A scale is a data collection tool designed to quantitatively assess the characteristics of research subjects such as behavior, attitude, observation, or theory.^{11,13} The scale development process is initiated according to the gap identified in the field. The literature states that scale development studies should first ensure that no scale related to the subject in question exists in the field.^{18,26} Ajzen states that scale development processes should be carried out according to the research to be conducted with TPB, the behavior/behavioral intention to be investigated, the population in which the research will be applied, and other factors in the research process.²⁷⁻³⁰ In our study, since it was determined that there was no valid and reliable measurement tool in the field that measured empathic ability and dispositional intention in line with the TPB, a scale was developed for use in the research. Among scaling techniques such as Thurstone-type and dimensional separation-type, Likert-type scoring is recommended for quantitative research because it

allows for numerical scoring and is suitable for evaluation.^{12,18,26}

For this reason, the Likert-type scaling technique, one of the scaling techniques, was used in the study. It is stated that a literature review should be conducted during the item pool creation process, possible sub-dimensions should be considered after the review, and three to four times the number of items anticipated for the scale should be prepared.^{18,29,30} As a result of the literature review conducted in our study, a 44-item pool covering the three components of TPB was created. The literature indicates that for the content validity stage, experts on the subject under investigation should first be identified, the item pool should be shared with the experts, and the necessary expert opinion on the items should be obtained.¹³ There are several techniques for evaluating expert opinions, one of which is the Lawshe technique.^{13,14} The Lawshe technique was used in the expert opinion phase of the study, and the item pool was presented to six experts who had worked on TPB, empathy, and scale

development. Regarding the Lawshe Technique, CVR index is calculated based on the number of experts and their evaluations of the items. According to the evaluation, if there are six experts, each item's index must be at least 0.99.¹³⁻¹⁵ In our study, the index score of each item remaining after the Lawshe Technique was 1.00.

Following expert review, it is recommended that a pilot study with a sample group of at least 30 individuals be conducted for the EFA assessment of the final draft scale form.^{11,18} In this context, a survey was administered to 30 midwives to ensure the comprehensibility and feasibility of the item pool created. Following the pilot study, it is reported that the EFA process should be initiated and the sample size should be increased to 5-10 times the number of items.^{11,17} In line with this information, 131 midwives were reached. The literature indicates that the KMO and Bartlett's Sphericity Test should be performed prior to EFA to determine the adequacy of the sample group and the relationship between the variables that constitute the scale. The KMO value must be at least 0.50, and the sphericity test must also be significant ($p < 0.05$).¹⁸⁻²⁰ KMO measure of sampling adequacy was 0.87, demonstrating that the sample was appropriate for factor analysis. The sphericity test results were also found to be significant ($p = 0.00$). When the obtained values are appropriate, the factor analysis stage is initiated. The literature suggests using principal components and rotation methods such as Varimax or Oblimin in the construct validity stage. As a result of the examinations, items with factor loadings below 0.30 must be removed from the scale.¹⁷⁻¹⁹ In our study, principal components and the Direct Oblimin Rotation method were used for factor analysis due to the relationship. After removing the items with a factor loading below 0.30 from the scale, a scale with a total of 15 items and three sub-dimensions was obtained. The Cronbach Alpha value of the developed scale is expected to be 0.70 or higher, while also accounting for 40-60% of the total variance.¹⁷⁻²⁰

The overall internal consistency of the scale was high, with a Cronbach's alpha coefficient of 0.89. The three subscales yielded Cronbach's alpha values of 0.90, 0.81, and 0.85, respectively. In addition, the three-factor structure accounted for 68.12% of the total variance. These findings indicate that the instrument demonstrates satisfactory reliability and internal consistency.

The literature suggests moving on to the CFA stage after EFA to evaluate factor relationships. At this stage, the scale's specific model fit indices are examined. The indices include "CMIN, DF, CMIN/DF, CFI, NNFI (TLI), RMSEA, SRMR" values.^{21-23,25} In our study, as a result of CFA analysis and the necessary covariances, the values were determined to be acceptable. Although the NFI (0.88) and RFI (0.85) values were slightly

below the conventionally recommended threshold of 0.90, these indices should not be interpreted in isolation. Previous methodological literature has emphasized that no single fit index is sufficient to determine model adequacy and that model fit should be evaluated using multiple complementary indices rather than relying on a single cutoff value. Furthermore, the performance of incremental fit indices such as NFI and RFI are influenced by sample size and model characteristics, and strict cutoff values may not be appropriate across all model structures.²¹ In the present study, the acceptable values obtained for the remaining goodness-of-fit indices collectively support the adequacy of the three-factor measurement model. Therefore, the slightly lower NFI and RFI values do not appear to compromise the overall model fit and are considered acceptable when interpreted alongside the other fit statistics.

As the final stage of the scale development process, it is recommended in the literature to apply the test-retest method at specific intervals and determine whether it shows consistency by examining the correlation coefficient.¹⁸ In our study, the scale was administered to 50 midwives at two-week intervals, and the values were found to be consistent.

CONCLUSION

The results of this study indicate that the Midwives' Empathetic Tendency and Skill Intention Scale (METSIS) is a valid and reliable measurement instrument. The scale provides a structured and theory-based approach to assessing midwives' empathetic tendencies and skill intentions within the framework of the TPB.

METSIS has the potential to serve as a valuable assessment tool in both clinical practice and research by enabling the systematic evaluation of empathetic intentions and skills among midwives. The instrument may serve as a valuable tool for identifying aspects that require further development and for guiding the design of tailored educational and behavior-oriented interventions. Moreover, future research is encouraged to examine midwives' empathetic skills and intentions using different methodological approaches and intervention designs, thereby contributing to the advancement of evidence-based midwifery practice and the enhancement of quality of care.

Limitations and Strengths

The findings of this study should be interpreted in light of several limitations. First, because the study population consisted exclusively of midwives from one country, the applicability of the results to different cultural settings and healthcare systems may be limited. Second, the use of self-report measures may have introduced social desirability bias, potentially

resulting in responses that do not fully reflect participants' actual empathetic behaviors in clinical practice.

Notwithstanding its limitations, the study demonstrates several important strengths. The scale was grounded in a well-established theoretical model, the TPB, which contributes to its conceptual soundness. This instrument was developed through a comprehensive methodological framework that includes expert review, exploratory and confirmatory, factor analyses, and test-retest reliability assessment, providing strong evidence for the psychometric robustness of the scale. Furthermore, the use of sufficiently large and independent samples for both EFA and CFA enhances the credibility of the identified factor structure. Notably, this study presents a culturally appropriate, valid, and reliable instrument specifically developed to assess midwives' empathetic skills and intention tendencies, thereby filling an important gap in midwifery literature and offering a useful resource for future research, educational settings, and clinical practice.

Ethical Considerations

An application was submitted to the Istanbul University-Cerrahpaşa Non-Interventional Clinical Research Ethics Committee for this study, and ethical committee approval was obtained on 21.02.2023 with the reference number E-74555795-050.01.04-626089.

Author Contributions

Study idea/design: TB, FA

Data collection: TB

Data analysis and interpretation: TB, FA

Literature review: TB

Writing of the article: TB, FA

Critical review: TB, FA

Final approval and accountability: TB, FA

Conflict of Interest: The authors declare no conflict of interest.

Financial Support: The authors declare no financial support.

REFERENCES

1. Hoffman ML. Empathy and Moral Development: Implications for Caring and Justice. Cambridge University Press, 2000.
2. Dökmen Ü. Communication Conflicts and Empathy. 66th ed. Remzi Bookstore, Istanbul, 2023.
3. Aktaş S, Pasinlioğlu T. The effect of empathy training given to midwives on meeting postpartum mothers' expectations and their perceptions of birth and the midwife. *Journal of Gynecology-Obstetrics and Neonatology*, 2017;14(2):60-5.
4. Aydın Kartal Y, Aker S. Investigation of the relationship between empathic tendencies and altruism levels of midwifery students. *Samsun Journal of Health Sciences*, 2020;5(2):97-102.
5. International Confederation of Midwives (ICM), Essential Competencies for Midwifery Practice. The Hague, 2024. <https://internationalmidwives.org/> Accessed date: March 20, 2025.
6. Schulz AA, Wirtz MA. Midwives' empathy and shared decision-making from women's perspective: Sensitivity of an assessment to compare quality of care in prenatal and obstetric care. *BMC Pregnancy and Childbirth*, 2022;22:717. <https://doi.org/10.1186/s12884-022-05041-y>
7. Ajzen I. From intentions to actions: A theory of planned behavior. In: J. Kuhl, J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* Springer, Heidelberg, 1985, pp. 11–39.
8. Şaşmaz Z, Tengilimoğlu E. A systematic review of studies conducted within the context of the theory of planned behavior: The DergiPark sample. *Journal of Selçuk Tourism and Informatics Research*, 2024;5:23-35.
9. Naskar ST, Lindahl JMM. Forty years of the theory of planned behavior: A bibliometric analysis (1985–2024). *Management Review Quarterly*. 2025;76:573–632. <https://doi.org/10.1007/s11301-025-00487-8>
10. Tanhan F, Kasap S, Hasan F, Ünsal F. Development of the Teacher Empathy Scale: A validity and reliability study. *Humanistic Perspective*, 2022;4(3):493–511. <https://doi.org/10.47793/hp.1049794>
11. DeVellis RF. Scale Development: Theory and Applications. 3rd ed. (Turkish translation). Nobel Publishing, Ankara. 2017.
12. Tavşancıl E. Measurement of Attitudes and Data Analysis with SPSS. Nobel Publishing, Ankara. 2010.
13. Yeşilyurt S, Çapraz C. A roadmap for content validity in scale development studies. *Erzincan University Journal of Education Faculty*, 2018;20(1):251-64. <https://doi.org/10.17556/erziefd.297741>
14. Lawshe CH. A quantitative approach to content validity. *Personnel Psychology*, 1975;28(4):563-75.
15. Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 2014;47(1):79–86. <https://doi.org/10.1177/0748175613513808>

16. Güldüren C, Çetinkaya L, Keser H. Development of an information security awareness scale for secondary school students. *Elementary Education Online*, 2016;15(2):682–95.
<https://doi.org/10.30703/cije.1131545>
17. Kishore K, Jaswal V, Kulkarni V, De D. Practical guidelines to develop and evaluate a questionnaire. *Indian Dermatology Online Journal*, 2021;12(2):266-275.
https://doi.org/10.4103/idoj.IDOJ_674_20
18. Büyüköztürk Ş. Handbook of Data Analysis for Social Sciences. 25th ed. Pegem Academy, Ankara. 2019.
19. Çokluk Ö, Şekercioğlu G, Büyüköztürk Ş. Multivariate Statistics for Social Sciences: SPSS and LISREL Applications. Pegem Academy, Ankara. 2010.
20. Field A. Discovering Statistics Using SPSS. 3rd ed. Sage Publications, London. 2009.
21. Byrne BM. Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming. 3rd ed. Routledge. 2016.
22. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. 8th ed. Cengage Learning, 2019.
23. Kline RB. Principles and Practice of Structural Equation Modeling. 4th ed. Guilford Press, 2015.
24. Tabachnick BG, Fidell LS. Using Multivariate Statistics. 7th ed. Pearson, 2019.
25. Schumacker RE, Lomax RG. A Beginner's Guide to Structural Equation Modeling. 4th ed. Routledge, 2016.
26. Bayat B. Measurement, scales and Likert scale development techniques in applied social science research. *Gazi University Journal of Economics and Administrative Sciences*, 2014;16(3):1–24.
27. Fishbein M, Ajzen I. Predicting and Changing Behavior: The Reasoned Action Approach. Psychology Press, New York, 2010.
28. Ajzen I. Attitudes, Personality and Behavior. 2nd ed. Open University Press, McGraw-Hill Education, 2005.
29. Ajzen I. Constructing a Theory of Planned Behavior questionnaire: Conceptual and methodological considerations, 2006 (Revised April 17, 2024). Available at: <https://people.umass.edu/aizen/faq.html> Accessed date: April 28, 2025).
30. Ajzen I. The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, 2020;2(4):314-24. <https://doi.org/10.1002/hbe2.195>