Original Article

Validation of the Persian version of the Ward Atmosphere Scale

Masoud Bahreini[1]; Mohammad Ali Abolhosseini[2]; *Shahnaz Pouladi[3]; Razieh Bagherzadeh[3]

[1]*Professor, Nursing and Midwifery School, Bushehr University of Medical Sciences, Bushehr, Iran.*
[2]*M.Sc, Instructor, Bushehr University of Medical Sciences, Bushehr, Iran.*
[3]*Associate professor, Nursing and Midwifery School, Bushehr University of Medical Sciences, Bushehr, Iran.*

Abstract

Introduction: Assessing the characteristics of the atmosphere of psychiatric wards provides managers with valuable information. This tool has had different accreditation features in different countries. This study was conducted to validate the Persian version of the short form of the psychiatry Ward Atmosphere Scale (WAS).

Materials and Methods: This cross-sectional study was part of a wider research (181 nurses and 299 patients) in selected hospitals of Bushehr and Shiraz cities-Iran in 2019. The psychiatry WAS-short form was validated according to the collected data. The steps of translation and re-translation of the scale were done by two people fluent in Farsi and English. Form validity and content validity were checked by asking the opinions of 11 experts, and the reliability of the questionnaire was checked by determining internal consistency and stability. SPSS software version 22 was used for data analysis.

Results: No items were removed during content validity. The average index and content validity ratio were 0.92 and 0.88, respectively. The Kuder-Richardson coefficient was calculated for nurses and patients, respectively, 0.92 and 0.86. The intracluster correlation between test and retest scores for nurses and patients was calculated as 0.91 and 0.89, respectively.

Conclusion: Considering the validity and reliability of the short form of the psychiatry ward atmosphere scale in Iranian society, it is recommended to be used by researchers and medical staff. Other methods, such as divergent validity and factor analysis of the questions of the revised Iranian version of the questionnaire, are suggested.

Keywords: Atmosphere, Psychiatry, Reliability, Validity, Ward

Introduction

The ward atmosphere comprises the physical setting, general structure, and social relations, undeniably affecting nurses' performance and patients' treatment outcomes (1). According to researchers, as the positive atmosphere changes the ward to a safe place for treatment and work, the negative atmosphere of treatment settings causes verbal and physical violence, increases serious injuries, and even patient death (2).

Psychiatric wards have a distinct atmosphere compared to other treatment-care wards (3). Confirming the above distinction in the past seventy years, the World Health Organization (WHO) has paid attention to the psychiatric ward atmosphere as an important and effective determinant of therapeutic measures and care interventions (4). According to the results of studies, ward atmosphere is correlated with some clinical outcomes, such as patient

*Corresponding Author:
Nursing and Midwifery School, Bushehr University of Medical Sciences, Bushehr, Iran.
pouladi2008@gmail.com

satisfaction, the health of treatment staff, personal independence (4), and human dignity (3). In the past half-century, numerous studies have evaluated the quality of care (5), psychosocial environment (6), physical environment (7), and risk factors (8) in psychiatric wards. Furthermore, researchers have sought to design and psychometrically evaluate the psychosocial atmosphere assessment tools in psychiatric wards as an undeniable necessity (5,9,10).

Moos designed and introduced the ward atmosphere scale (WAS) in 1968 (11,12). This tool measures the psychiatric ward atmosphere in the form of long (100 items) and short (40 items). Most studies have focused on the psychometrics of the 100-item form of this tool (4). Even though the results of some studies indicate that WAS subscales have an acceptable internal structure, others report that some subscales of this questionnaire do not have desirable reliability (12-14). Given that past studies have reported instabilities and differences in revising the original version of WAS in different societies, the Persian version of this scale has yet to be psychometrically evaluated in Iran. The evaluation of the validity and reliability of this tool can be taken into consideration in Iran. Most previous studies evaluated the validity and reliability of the ward atmosphere scale- long form with 100 items (15). The short form of this tool seems more acceptable to patients of psychiatric wards and other communities owing to the fewer items. Given the increasing trend of psychiatric diseases and challenges in psychiatric wards, the importance of evaluating the psychiatric ward atmosphere, its effects on the treatment process, and paying special attention to the fact that the atmosphere of each psychiatric center has its characteristics, the present study aimed to translate, culturally adapt, and validate the ward atmosphere scale- short form in certain hospitals in Bushehr and Shiraz in 2019.

Materials and Methods

The present cross-sectional study was taken from a wider research (181 nurses/ paramedics and 299 patients) to validate the psychiatry ward atmosphere scale-short form in certain hospitals in Bushehr and Shiraz in 2019. The statistical population consisted of all nurses and paramedics and hospitalized patients in Ibn Sina and Moharreri psychiatric hospitals in Shiraz,

the psychiatric wards of Shohadaye-Khalij-e-Fars Hospital in Bushehr, and Shahid Ganji Hospital in Borazjan. The equation of sample size for regression was utilized to analyze the data of nurses and paramedics in the present research. The sample size of nurses was obtained equal to 140 based on 14 variables, including 10 samples for each main variable, and considering all demographic variables as effective factors. In these hospitals, 181 nurses and paramedics had the inclusion criteria. Finally, all nurses working in these centers (i.e., 181 nurses and paramedics) were included in the research. Considering three main variables, and all demographic variables as possible effective factors, 13 variables were considered for patients.

Further, 130 samples were obtained for patients based on the sample size formula for regression, and finally, 156 samples were obtained considering a 20% drop. In the centers above, 299 patients had the inclusion criteria, and thus, all were included in the study. The inclusion criteria for nurses and paramedics were consent to participate in the study, working in the psychiatric ward, and experience of at least six months of working there. The inclusion criteria for patients were the consent of the patient and family to participate in the study at least seven days after admission to the neuropsychiatric ward and the patient's ability to answer questions in writing or orally. The exclusion criteria for patients and nurses were the psychiatric emergency for the patient and incomplete questionnaires by the group of nurses, paramedics, and patients. The convenience sampling method was conducted. To this end, all nurses and paramedics in the morning, evening, and night shifts and patients of the centers were included in the study. The measurement tool was a 40-item Moos psychiatric ward atmosphere scale-short form validated in this study. This tool comprised ten subscales, namely "participation in treatment", "support", "spontaneous behavior", "independence", "operational familiarity", "familiarity with personal problems", "aggression", "order and discipline", "clarity of the treatment plan", and "control by personnel". Each subscale contained four items. Answers are given as true or false. The measurement steps constituted the translation and re-translation, pre-test and cognitive interview, evaluation of the content validity, quantitative and qualitative face validity, determination of reliability, and test stability.

The first stage of validation

Translating the ward atmosphere scale of psychiatric hospitals comprised the translation, integration, and comparison of the translation, returning the final translated version from the target to the original language, pre-test, and cognitive interviews. In the translation step, the original version was translated from English to Farsi. First, the original English version was translated into Farsi by two translators fluent in English and Farsi. In the step of integration and comparison of the translation, a meeting was held with four people, including two experts in instrumentation, a Ph.D. in psychiatric nursing, and a post-graduate nursing student with clinical experience. The contradictions between the original version and the translation of the questionnaire were resolved, and the initial version was created. Thereafter, the final version of the questionnaire was given to two people fluent in Farsi and English, who did not participate in the previous steps, for translating the final version from the target to the original language. The group of four re-checked the translated version, and the contradictions were resolved. In the pre-test and cognitive interview step, 20 patients and 20 nurses/paramedics were asked about their understanding of the items (Do they have to repeat the questions or items in their languages to understand the questions, and what comes to their mind when they hear a certain phrase of the questionnaire). The final version was prepared after the pre-test.

The second stage of validation

Face validity and content validity were used qualitatively and quantitatively to determine the validity of the questionnaire. Face-to-face interviews were conducted with 20 nurses and 20 patients for qualitative face validity. They were asked about the level of difficulty, proportionality, ambiguity of the items, and the need to remove or integrate them. Then, the items were revised according to the recommendations of this group. Quantitative face validity was also examined by measuring the impact scores of the items. To perform quantitative face validity, ten patients and nurses were asked to evaluate the items in terms of the importance of each item and give each item a score of 1 to 5 based on the importance. A score of 1 indicated the minimum, and a score of 5 indicated the maximum importance of the item. The impact score for each item was calculated by multiplying the importance of an item by the number of its repetitions. A score higher than 1.5 was considered appropriate (16). Content validity was also evaluated qualitatively and quantitatively. Content validity refers to the extent to which the items of the tools are related to the research content or dimensions of the concept (17). In the qualitative determination of content validity, ten experts in the psychiatric ward atmosphere were asked to evaluate the questionnaire items in terms of grammar, use of suitable words, and placement of the items in their right places. The Content Validity Ratio (CVR) and Content Validity Index (CVI) were used to determine content validity quantitatively. To determine CVI, 11 experts, including four instrument specialists, three psychiatric nursing doctors, two psychiatrists, one health psychology doctor, three psychiatric nursing masters with more than two years of coaching experience in a psychiatric hospital, and one psychologist with nine years of work experience in the psychiatric ward, and a head nurse of the psychiatric ward, were asked to answer the items of all three questionnaires regarding three criteria, simplicity, clarity, and relevance, on a 4-point scale. After that, the validity index for each of the three criteria was calculated by dividing the number of experts who gave each criterion a score of 3 and 4 (e.g., the number of those who gave scores to completely relevant and acceptable relevance regarding the relevance criterion), by the total number of experts who gave their opinion.

A score of 0.79 and above was considered suitable. The CVI was measured separately for each item, and its mean was calculated for the whole questionnaire. According to Lawshe, the CVR determines the necessity of an item in the instrument on a three-point Likert scale (necessary, useful but unnecessary, and unnecessary). CVR was calculated using the CVR strict and CVR relaxed methods. In CVR strict, the option "necessary" receives a score of 1, and the other two options are given a score of zero (18). In the CVR relaxed method, the necessary option receives a score of 2, and the "useful but unnecessary" option is given a score of 1. It is then calculated according to the following equation.

$$CVR = \frac{\text{Number of necessary responses of each item} - \frac{\text{Total number of participants}}{2}}{\frac{\text{Total number of participants}}{2}}$$

Based on the Lawshe table, the items with ratios greater than 0.59 (according to the evaluation of 11 professors) were retained to determine the minimum value of CVR. In other words, the existence of an item with an acceptable level of significance ($P < 0.05$) is necessary and important in this tool (19). The CVR was calculated separately for each item of the questionnaire. Further, its mean was measured for the whole questionnaire. It was decided to retain or delete the items according to all three criteria (item impact, CVR, and CVI). Its internal consistency and stability were evaluated to determine reliability in such a way that a pilot study was conducted on nurses and patients in the psychiatric ward. Therefore, 20 patients and 20 nurses were participated in the study. The reliability was evaluated by determining internal consistency and stability. The Kuder-Richardson (KR) coefficient was utilized to determine the internal homogeneity due to the bimodality of the response. KR coefficient above 0.7 was considered satisfactory internal consistency.

The stability evaluation was performed using the test-retest method. The interval between the test and re-test was two weeks. After collecting both stages of data, the Intra-class Correlation Coefficient (ICC) (20) and the type of consistency with the random-bivariate model were calculated for subscales and all questionnaires. Further, the confidence interval and significance level were measured for ICC. If ICC is 0.8 or higher between two tests, it indicates satisfactory stability (21,22). As discriminant validity, the correlation between the questionnaire subscales was compared with the internal correlation of each subscale. Therefore, the greater internal correlation of the subscale than the correlation of each subscale with other subscales and the correlation of the subscale with the whole questionnaire indicated the appropriate discriminant validity.

The legal permits were obtained from the Research and Technology Deputy of Bushehr University of Medical Sciences for compliance with ethical issues and were approved with an ethical code IR.BPUMS.REC.1398.080 by the Ethics Committee. All participants signed the informed consent forms and received the necessary explanations about completing the questionnaires and the research objectives. The participants' voluntary and anonymous participation and confidentiality were also observed.

Results
The nurses' mean age and work experience were 36.7 and 10.42 years respectively. Further, 59% of them were male, and the rest were female. More than 70% of the nurses were official employees, and more than 90% were rotational shift nurses. The nurses' mean working hours per month was 174.96 hours. More than 75% of the patients were males, and most were self-employed or unemployed. Furthermore, most patients had secondary school diplomas or lower, and they often needed to be covered by supporting organizations (Tables 1 and 2).

Table 1. The demographic quantitative variables related to nurses and patients

Patients	Variable	Mean ±SD	Median	Mode
	Age	31.53 ± 6.17	-	-
	Number of hospitalization day	21.22±7.22	19.00	18.00
	Number of children	0.871 ± .07	0.00	0.00
Nurses	Age (Year)	36.7 ± 7	-	-
	Work experience (Total) (Year)	10.42± 5.48	-	-
	Work experience in current ward (Year)	5.80± 3.77	-	-
	Number of night shifts per month	4.76 ±2.10	5	4
	Working hours per month	174.96± 43.47	-	-
	Interest in the job	5.91± 2.68	6	5

Table 2. Frequency of qualitative demographic characteristics of nurses and patients

Variable	Variable classes	Number	Percentage
Gender	Male	231	77.3
	Female	68	22.7
Marital Status	Single	142	47.6
	Married	131	44.3
	Divorced	24	8.1
Job	Unemployed	121	40.7
	Manual worker	15	5.1
	Freelance job	127	42.8
	Employee	34	11.4
Education	High school	106	35.5
	Diploma	149	49.8
	Above the diploma	44	14.7
Coverage of support organization	Yes	66	22.1
	No	232	77.9
Place of hospitalization	Bushehr and Borazjan	25	8.4
	Ibn Sina	90	30.1
	Moharreri	184	61.5
Gender(nurses)	Male	108	59.1
	Female	73	40.3
Marital Status(nurses)	Single	40	22.1
	Marriage	141	77.9
Hospital of service(nurses)	Bushehr and Borazjan	23	12.7
	Ibn Sina	62	34.3
	Moharreri	96	53.0
Employment status(nurses)	Official	133	73.5
	Training course	21	11.6
	Other	27	14.9
Shift work(nurses)	Rotating shift	168	92.8
	Fix shift	13	7.2

In qualitative face validity, the way of writing two items was edited based on the patients' opinions to make it easier for people to understand. In the process of quantitative face validity, the results obtained from calculating the item impact score indicated that the item impact score was greater than 1.5 for all items (the minimum item impact was 4.14, and the maximum was 5). Therefore, all items were suitable for evaluating content validity. Minor changes were made in the way the items were written according to the evaluation of qualitative content validity. For example, the written format was as follows in item 20: "Patients can interrupt the doctor when he is speaking". It was changed according to the content review group's opinion and the scoring of the questionnaire: "Patients may interrupt the doctor when he is speaking". The mean validity index obtained from the three criteria (simplicity, clarity, and relevance) was above 0.79 for all items except for item 23, which was equal to 0.75. This item was not removed due to the appropriateness of the item's impact and content validity ratio. The mean content validity index was 0.87 for the whole instrument. The content validity ratio was greater than the cut-off point for all items, i.e., greater than 0.59. The mean CVR strict and CVR relax were 0.88 and 0.99 respectively (Table 3).

The internal consistency of the questionnaire was 0.92 in nurses and 0.86 in patients using the Kuder-Richardson coefficient. The stability of the questionnaire was 0.91 and 0.86 respectively for nurses and patients using the ICC of test-retest (Tables 4 and 5). The correlation between each subscale and other subscales was also calculated. All cases showed a correlation of two subscales lower than the internal correlation of both subscales (less than 0.6 in all cases).

Table 3. Values and average CVI, CVR, and impact score statements of Ward Atmosphere Scale

Number of statement	Statements	CVI				CVR		Impact score
		Relevance score	Clarity score	Be simple score	Average indicators	CVR strict	CVR relax	
1	Patients put a lot of energy into what they do around here	0.91	0.91	0.82	0.88	0.64	1	5
2	Doctors have very little time to encourage patients	1	0.91	1	0.97	1	1	5
3	Patients tend to hide their feelings from one another	0.91	0.91	0.91	0.91	1	1	5
4	The staff act on patients' suggestions	0.82	0.82	1	0.88	0.82	1	5
5	New treatment approaches are often tried in this program	0.91	0.82	0.82	0.85	0.82	1	4.3
6	Patients hardly ever discuss their sex life	1	0.91	0.91	0.94	0.82	1	4.3
7	Patients often gripe	0.91	0.91	0.91	0.91	0.64	1	4.3
8	Patients activities are carefully planned	1	1	1	1	0.82	1	4.9
9	The patients know when doctors will be on the unit	1	0.91	0.91	0.94	1	1	5
10	The staff very rarely punish patients by restricting them	1	0.91	0.91	0.94	0.82	1	4.9
11	This is a lively program	1	0.91	1	0.97	0.64	1	4.8
12	The staff know what the patients want	0.91	0.91	0.91	0.91	1	1	5
13	Patients say anything they want to the doctors	1	0.72	0.91	0.88	1	1	5
14	Very few patients have any responsibility here	1	0.82	0.91	0.91	1	1	4.3
15	There is little emphasis on teaching patients solutions to practical problems	1	0.91	0.91	0.94	1	1	5
16	Patients tell each other about their personal problems	1	1	1	1	1	1	4.3
17	Patients often criticize or joke about the staff	0.82	0.82	0.82	0.82	0.64	0.82	4.3
18	This is a ve1y well organized program	0.91	0.82	0.72	0.82	0.82	1	4.3
19	Doctors do not explain what treatment is about to patients	0.91	0.82	1	0.91	1	1	5
20	Patients may interrupt when a doctor is talking	1	0.91	1	0.97	0.82	1	4.9
21	The patients are proud of this program	0.91	0.91	1	0.94	0.82	1	4.9
22	Staff are interested in following up patients when they discharge	1	1	0.91	0.97	1	1	5
23	It is hard to tell how patients are feeling here	0.82	0.72	0.72	0.75	0.82	1	4.9
24	Patients are expected to take leadership here	0.91	1	1	0.97	1	1	4.9
25	Patients are strongly encouraged to plan for the future	1	1	1	1	1	1	5
26	Personal problems are openly talked about	0.91	0.91	0.91	0.91	1	1	5
27	Patients in this program rarely argue	1	1	1	1	1	1	4.9
28	The staff make sure that the unit is always neat	0.91	0.82	0.72	0.82	0.82	1	5
29	A nurse or doctor will always explains why patient's medicine is changed	1	1	1	1	1	1	5
30	Patients who break the rules are punished for it	1	1	1	1	0.82	1	5
31	There is very little group spirit in this program	1	1	1	1	0.82	1	3.7
32	Nurse has Very little time to encourage patients	1	1	0.82	0.94	0.82	0.82	5
33	Patients are careful about what they say when staff are around	1	0.71	0.91	0.87	1	1	5
34	Patients here are encouraged to be independent	1	0.91	0.91	0.94	1	1	5
35	There is little emphasis on what patients will be doing after discharge	0.91	0.82	0.72	0.82	0.64	1	4.9
36	Patients are expected to share their personal problems with each other	1	1	1	1	0.82	1	5
37	Staff sometimes argue openly with each other	0.91	0.91	0.91	0.91	0.64	1	5
38	The unit sometimes gets very messy	0.91	1	0.91	0.94	1	1	5
39	Patients who argue with other patients will get into trouble with the staff	1	1	1	1	1	1	5
40	The patients clearly understand the program rules	0.91	0.91	0.91	0.91	0.82	1	4.9
	Average	0.95	0.91	0.92		0.88	0.99	4.8

Table 4. Result intra correlation and consistency for Ward Atmosphere Scale in nurses

Subscale	Coefficient Kuder-Richardson for subscale	ICC	Confidence interval for ICC	P value for ICC	Correlation subscale with total score
Involvement	0.841	0.974	0.942-0.992	<0.001	0.650
Support	0.812	0.921	0.878-0.954	<0.001	0.624
Spontaneous behavior	0.721	0.951	0.904-0.981	<0.001	0.654
Autonomy	0.711	0.908	0.826-0.974	<0.001	0.648
Practical orientation	0.770	0.899	0.746-0.960	<0.001	0.510
Personal problems Orientation	0.840	0.949	0.878-0.974	<0.001	0.678
Anger and aggression	0.898	0.962	0.961-0.982	<0.001	0.631
Order and organization	0.868	0.928	0.819-0.965	<0.001	0.624
Program clarity	0.873	0.958	0.892-0.988	<0.001	0.688
Staff control	0.830	0.986	0.955-0.995	<0.001	0.652
Total score of the ward atmosphere	0.922	0.912	0.777-0.961	<0.001	1

Table 5. Result intra correlation and consistency for Ward Atmosphere Scale in patients

Subscale	Coefficient Kuder-Richardson for subscale	ICC	Confidence interval for ICC	P value for ICC	Correlation subscale with total score
Involvement	0.798	0.914	0.912-0.950	<0.001	0.622
Support	0.812	0.887	0.858-0.934	<0.001	0.601
Spontaneous behavior	0.721	0.851	0.811-0.881	<0.001	0.618
Autonomy	0.740	0.828	0.736-0.928	<0.001	0.638
Practical orientation	0.770	0.899	0.856-0.950	<0.001	0.651
Personal problems Orientation	0.869	0.849	0.808-0.874	<0.001	0.523
Anger and aggression	0.884	0.902	0.861-0.963	<0.001	0.604
Order and organization	0.743	0.918	0.880-0.965	<0.001	0.650
Program clarity	0.771	0.891	0.832-0.908	<0.001	
Staff control	0.890	0.886	0.885-0.915	<0.001	0.581
Total score of the ward atmosphere	0.867	0.892	0.855-0.948	<0.001	1

Discussion

The present study aimed to validate the Persian version of the psychiatry Ward Atmosphere Scale. The results of CVI, CVR, as well as the Kuder-Richardson coefficient were high for all subscales of the questionnaire in nurses and patients, indicating the desirable validity and reliability of the questionnaire.

The present study utilized a Ward Atmosphere Scale-short form (WAS), while there was no article that psychometrically evaluated this version in domestic and foreign studies. Sorlie and Rossberg reported that the removal of 32 items and 16 items, respectively, from a 100-item version of psychiatric WAS increased the validity and reliability of the revised questionnaire (14,23). The present

results confirmed the research by Rossberg et al. However, Rossberg revised the range of answering the questionnaire questions on a four-point Likert scale (totally agree= 3 to disagree= 0) with the idea that patients' ways of answering the questions on a true/false scale limited responses. Furthermore, he stated that removing some items and developing the response scale increased the psychometric results in the Norwegian version of WAS. However, the present study reported that answering questions in a wide range required high concentration and accuracy and was worth considering in this group of patients regarding cognitive and psychological problems. Therefore, using the same main scale of WAS in true/false levels was preferred. Confirming

this issue, Bakken studied the reliability of WAS in patients with mild and moderate mental disorders and reported that only patients with mild disorders could answer the questions of WAS without receiving the help of others, while patients with moderate mental disorders could not answer more than 50% of questions without any help, indicating the importance of the necessary accuracy in answering the items (24). In the present study, reliability was measured in both patients and nurses. Based on the results, there needed to be a higher reliability of the questionnaire among nurses and patients. The results were consistent with a study by Schröder et al. on the reliability of the Quality of Psychiatric Care-Inpatient Staff (QPC-IPS) questionnaire in nurses and the quality of care for psychiatric patients (5). The small difference in internal consistency and the intra-class correlation coefficient between patients and nurses in the present study was because patients in the acute stage of the disease were not included in the pilot study. Furthermore, the questionnaire had desirable reliability, at least for psychiatric patients who were not acute. The evaluation of reliability, stability, and other validation criteria of the questionnaire between groups of patients with different severities of psychiatric diseases can create more accurate results and indicate the usability of this questionnaire for acute patients.

The results of the present study were consistent with a study that aimed to determine the psychometrics of the Swedish version of WAS that indicated the desirable validity and reliability of the revised scale (13). It is worth noting that even though a great number of items in a questionnaire can increase the internal consistency of the questionnaire, and internal consistency is greatly dependent on the increase in the number of items, the consistency and stability in the present study are higher than studies that psychometrically evaluated a 100-item WAS (14,23), indicating that a large number of items decreases accuracy and inconsistent response in the target population, and thus decreases reliability. Wann-Hansson and Eklund measured the patients' general

opinions on the psychiatric ward atmosphere using the Good Milieu Index (GMI) and reported a low correlation between the WAS and GMI. This issue challenges the construct validity of WAS (13).

Like other studies, the present study had some strengths and limitations. The initial validation of the Moos psychiatric ward atmosphere scale in Iran was the strength of the present study among psychometric studies. An important research limitation was the need for construct validity. Future studies are suggested to utilize methods such as the evaluation of divergent validity and factor analysis of questionnaire items, as well as the correlation between WAS and GMI, to strengthen the psychometric results of the revised Iranian scale. Given that this scale was validated in Iranian culture, the results cannot be generalized to other cultures due to the different psychometric results in different countries. The present research used a pilot study with a small sample size to determine reliability and stability. Repeating the validation with a larger sample size can provide more accurate results.

Conclusion

The results of the present study indicated that the psychiatric Ward Atmosphere Scale-short form (WAS) in the Iranian population had desirable validation characteristics. Researchers in Iran can use this questionnaire. The Persian version of the psychiatric WAS-short form can be a basis for further studies in this field because it was first validated in Iran.

Acknowledgments

References

1. Fröhlich D, Rabenschlag F, Schoppmann S, Borgwardt S, Lang UE, Huber CG. Positive effects of an anti-aggression and de-escalation training on ward atmosphere and subjective safety may depend on previous training experience. Front Psychiatry 2018; 9: 134-8.
2. Baumgardt J, Jäckel D, Helber-Böhlen H, Morgenstern K, Voigt A, Schöppe E, et al. Making psychiatric wards more peaceful places: Evaluating the implementation of the safewards model in acute psychiatry using a pre-post mixed-method study design. California: Sage; 2020.

3. Weltens I, Drukker M, van Amelsvoort T, Bak M. Staff and ward factors associated with aggression development on an acute closed psychiatric ward: an experience sampling method study. BMJ Open 2023; 13(2): e067943.

4. Dickens GL, Johnson A, Steel K, Everett B, Tonkin M. Interventions to improve social climate in acute mental health inpatient settings: Systematic review of content and outcomes. SAGE Open Nurs 2022; 8: 23779608221124291.

5. Schröder A, Lundqvist LO. The quality in psychiatric care-inpatient staff instrument: A psychometric evaluation. Healthcare (Basel) 2022; 10(7): 1213.

6. Corbetta M, Corso B, Camuccio CA. Rules and ward climate in acute psychiatric setting: Comparison of staff and patient perceptions. Int J Ment Health Nurs 2022; 31(3): 611-24.

7. Belsiyal CX, Rentala S, Das A. Use of therapeutic milieu interventions in a psychiatric setting: A systematic review. J Educ Health Promot 2022; 11: 234.

8. Weltens I, Bak M, Verhagen S, Vandenberk E, Domen P, van Amelsvoort T, et al. Aggression on the psychiatric ward: Prevalence and risk factors. A systematic review of the literature. PLoS One 2021; 16(10): e0258346.

9. Tomlin J, Tonkin M. The EssenCES measure of ward atmosphere: Mokken scaling, confirmatory factor analysis, and investigating patient-level characteristics. Int J Forensic Ment Health 2023; 22(3): 199-209.

10. Vahidi M, Namdar Areshtanab H, Ebrahimi H, Asghari Jafarabadi M. Development of "The Safe Psychiatric Ward Battery". Clin Nurs Res. 2023; 32(2): 375-83.

11. Alexiou E, Degl'Innocenti A, Kullgren A, Wijk H. The impact of facility relocation on patients' perceptions of ward atmosphere and quality of received forensic psychiatric care. J Forensic Leg Med 2016; 42: 1-7.

12. Efkemann SA, Bernard J, Kalagi J, Otte I, Ueberberg B, Assion H-J, et al. Ward atmosphere and patient satisfaction in psychiatric hospitals with different ward settings and door policies. Results from a mixed methods study. Front Psychiatry 2019; 10: 576-87.

13. Wann-Hansson C, Eklund M. A revised Swedish version of the Ward Atmosphere Scale: Usability and psychometrics. Nord J Psychiatry 2010; 64(5): 303-9.

14. Sørlie T, Parniakov A, Rezvy G, Ponomarev O. Psychometric evaluation of the Ward Atmosphere Scale in a Russian psychiatric hospital. Nord J Psychiatry 2010; 64(6): 377-83.

15. Banks C, Priebe S. Scales for assessing the therapeutic milieu in psychiatric inpatient settings: A systematic review. Gen Hosp Psychiatry 2020;66:44-50.

16. Deodato J. Evaluating web-scale discovery services: A step-by-step guide. Information technology and libraries 2015; 34(2): 19-75.

17. Yusoff MSB. ABC of content validation and content validity index calculation. Education in medicine journal 2019; 11(2): 49-54.

18. Betschart P, Staubli SE, Zumstein V, Babst C, Sauter R, Schmid H-P, et al. Improving patient education materials: A practical algorithm from development to validation. Curr Urol 2019; 13(2): 64-9.

19. Taghizadeh Z, Ebadi A, Montazeri A, Shahvari Z, Tavousi M, Bagherzadeh R. [Psychometric properties of health related measures. Part 1: Translation, development, and content and face validity]. Payesh (Health monitor) 2017; 16(3): 343-57. (Persian)

20. Shrotryia VK, Dhanda U. Content validity of assessment instrument for employee engagement. SAGE Open 2019; 9(1): 2158244018821751.

21. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. Archives of orofacial science 2017; 12(1): 1-11.

22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016; 15(2): 155-63.

23. Røssberg JI, Friis S. A suggested revision of the Ward Atmosphere Scale. Acta Psychiatr Scand 2003; 108(5): 374-80.

24. Bakken TL, Røssberg JI, Friis S. The ward atmosphere scale for psychiatric inpatients with intellectual disability: A pilot study. Adv Ment Health Intellect Disabil 2012; 6(5): 265-72.