Received: 24 May 2020 | Revised: 27 June 2020 | Accepted: 23 July 2020

DOI: 10.1002/nop2.609

RESEARCH ARTICLE

NursingOpen Open Access WILEY

Translation, reliability and validity of Japanese version the TeamSTEPPS® teamwork perceptions questionnaire

Takeshi Unoki[1] | Yujiro Matsuishi[2] | Takahiro Tsujimoto[3] | Ryuichi Yotsumoto[4] | Toru Yamada[4] | Yuka Komatsu[5] | Daisaku Kashiwakura[1] | Norimasa Yamamoto[6]

[1]Department of Adult Health Nursing, School of Nursing, Sapporo City University, Sapporo, Japan

[2]Pediatric Intensive Care Unit, University of Tsukuba Hospital, Tsukuba, Japan

[3]Department of Nursing, Nara Medical University Hospital, Kashihara, Japan

[4]Department of Nursing, Toho University Omori Medical Center, Tokyo, Japan

[5]Stroke Care Unit, Kyorin University Hospital, Mitaka, Japan

[6]Master's Course, Graduate School of Nursing, Sapporo City University, Sapporo, Japan

Correspondence
Takeshi Unoki, Department of Adult Health Nursing, School of Nursing, Sapporo City University, 13-1 Kita 11 Nishi, Chuou-ku, Sapporo 060-0011, Hokkaido, Japan.
Email: t.unoki@scu.ac.jp

Funding information
This work was supported by a 2018-2020 research grant from Sapporo City University.

Abstract

Aim: This study aimed to translate the TeamSTEPPS® Teamwork Perceptions Questionnaire (T-TPQ) into Japanese and assess its validity and reliability.

Design: Translation of the T-TPQ and a cross-sectional survey.

Methods: Following a forward and back translation of the questionnaire, content validity was assessed by an expert panel using item-level content validity index. Construct validity was assessed by a confirmatory factor analysis. Further, intraclass correlation coefficient was estimated by test–retest methods.

Results: A total of 587 healthcare professionals responded to the translated T-TPQ. The item-level content validity index ranged between 0.8 and 1.0, indicating an acceptable content validity. The multiple fit indices showed an acceptable fitting model. Fifty-one healthcare professionals participated in the test–retest method. Intraclass correlation coefficients for all dimensions ranged from 0.838 to 0.957, indicating acceptable test–retest reliability. Our findings suggest that the Japanese version of the T-TPQ has acceptable validity and reliability.

KEYWORDS
healthcare worker, quality and safety, reliability and validity, teamwork

1 | INTRODUCTION

Teamwork is essential for delivering safe and high-quality care in the current complex medical systems. Although studies have emphasized the importance of teamwork, barriers, such as silo mentality and hierarchies, inhibit teamwork amongst medical professionals (Weller, Boyd, & Cumin, 2014). The Department of Defense and the Agency of Healthcare Research and Quality (AHRQ) developed the Team Strategies and Tools to enhance Performance and Patient Safety (TeamSTEPPS®) approach to facilitate and integrate teamwork in clinical practice.

2 | BACKGROUND

Thus far, teamwork in medical settings has been emphasized in Japan; however, it was reported that teamwork within the hospital was still inferior to those in the United States according to a survey in 2013 (Fujita et al., 2013). Therefore, improvement of teamwork in the unit and hospital level was required in Japan. To investigate the level at teamwork accurately, teamwork measurement should be used in different languages and focus on teamwork at the unit or hospital level are preferred. Few Japanese international scales are available that measure teamwork at the unit level. The Hospital Survey on Patient

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2020 The Authors. Nursing Open published by John Wiley & Sons Ltd

Safety Culture developed by AHRQ, that has been internationally used, was translated in Japanese (Ito et al., 2011); however, it focused on safety and teamwork was just part of the scale.

The TeamSTEPPS® Team Perception Questionnaire (T-TPQ) (Battles, n.d.) was designed to assess an individual's perception of group- or unit-level teamwork skills and behaviour. It consists of 35 items that are equally distributed across five teamwork dimensions: Team Structure, Leadership, Situation Monitoring, Mutual Support and Communication. Each item is rated on a five-point Likert scale, ranging from 1 (strongly disagree with the statement)–5 (strongly agree with the statement). Sum or averages for each dimension of the T-TPQ are used to assess teamwork perception in a team. The T-TPQ has been validated (Keebler et al., 2014) and widely used across a range of settings, from quality improvement assessment in clinical practice to research studies. A recent review suggests that implementation of the TeamSTEPPS® is associated with improved communication, reduced rate of clinical errors and higher patient satisfaction (Parker, Forsythe, & Kohlmorgen, 2019). To facilitate teamwork, measurement of recognition of current teamwork status is necessary. Previous studies assessed and confirmed the construct validity and internal consistency of the original instrument (Keebler et al., 2014). The T-TPQ has been translated into Norwegian (Ballangrud, Husebø, & Hall-Lord, 2017) and Korean (Hwang & Ahn, 2015) only. Thus, the aim of this study was to translate the T-TPQ into the Japanese language and evaluate its validity and reliability.

3 | THE STUDY

3.1 | Design

The original T-TPQ was translated into Japanese, and its content validity, construct validity, test–retest reliability and internal consistency were evaluated. In these processes, we conducted several cross-sectional questionnaire surveys on healthcare professionals who were native Japanese speakers and worked in Japan.

3.2 | Method

3.2.1 | Translation process

Permission to translate the original instrument was obtained from the AHRQ. The translation process followed the back translation method based on a guideline (Sousa & Rojjanasrirat, 2011). A part of the translation process in this study was reported in the previous study (Unoki et al., 2020). The translation team consisted of two nurse educators and five clinical nurses. Our translation process was conducted in five steps, which included assessment of content validity.

STEP 1: Forward translation

Forward translation of the T-TPQ from English to Japanese was performed by two nurse educators. The nurse educators independently translated from English into Japanese. The two translations were synthesized into one preliminary version through discussion between both translators. Thereafter, the translation team assessed the equivalence of meaning between the original questionnaire and the translated version. In addition, the translation team critically appraised whether each version was easy to understand for the medical staff working in Japan.

STEP 2: Back translation

One native-English professional translator with adequate knowledge of healthcare terminology, and blinded to the original T-TPQ, back translated the T-TPQ from Japanese to English. The other native-English professional translator compared the original and back-translated versions and assessed the accuracy of terminology and clarity of expressions in terms of consistency in meaning. The assessment revealed a few inconsistencies in meaning between both versions.

STEP 3: First pilot test

A pilot test of the Japanese version of the T-TPQ was conducted to evaluate the clarity of items. A convenience sample of 43 healthcare professionals, including 32 nurses, 3 clinical engineers, 3 physical therapists, 2 medical check engineers and 2 physicians rated each item of the questionnaire on a dichotomous scale ("clear" vs. "unclear") via a web-based survey. We calculated the proportion of "unclear" response for each item. If participants rated an item as "unclear" or had suggestions for improvement, they were asked to comment on those items. Items rated as "unclear" by over 20% of the participants were revised by the translation team based on participants' feedback.

STEP 4: Evaluation of clarity and relevant validity by the expert panel

Item 3 of the "Team Function" dimension was rated as "unclear" by 21% of the participants, and thus, the translation team revised it based on this feedback.

The expert panel, which consisted of nine nurses, a physical therapist and a physician with knowledge of medical safety, assessed the clarity of the T-TPQ following the same process as in STEP 3. Items rated as "unclear" by over 20% of the expert panel were revised by the translation team. Item 7 of the "Team Function" dimension was rated as "unclear" by 27.3% of the expert panel, and the item was revised by the translation team based on the panel feedback. In addition, suggestions from the panel led to minor revision of each item. Care was taken to ensure that equivalence of meaning was not lost in these revisions.

Thereafter, the expert panel evaluated the content validity of all items. Each item was rated on a four-point Likert scale, where 1 = not relevant; 2 = unable to assess relevance; 3 = relevant but needs minor alteration; and 4 = very relevant and succinct. Content validity index (CVI) was determined by calculating the proportion of items with a rating of 3 or 4 given by the expert panel (Lawshe, 1975). The CVI of all items ranged between 0.8 and 1.0, thereby confirming the content validity of the questionnaire.

STEP 5: Second pilot test (evaluation of clarity)

A pilot test to assess the clarity of the items of the pre-final version was conducted on a convenience sample of 44 healthcare professionals including 33 nurses, 4 clinical engineers, 3 physical therapists, 2 physicians and 2 clinical check engineers. The process outlined in STEP 3 was again followed, and we confirmed that none of the items had over 20% "unclear" response. Thus, the Japanese version of T-TPQ was finalized (Supporting Information 1).

3.2.2 | Construct validity, internal consistency and test–retest reliability of the Japanese version of the T-TPQ

Participants

Participant recruitment for the evaluation of construct validity and internal consistency was carried out in two processes. First, we conducted a cross-sectional survey between February and March 2019 (Test 1). Nurses and midwives in a university hospital in Japan who provided direct patient care were selected, regardless of the department, unit and ward. The anonymous questionnaires included questions on participants' characteristics, such as gender, years of experience, working unit and certification, as well as the T-TPQ. The questionnaire was distributed to 876 participants by a nurse manager.

In the second process (Test 2), we attempted to collect data from medical professionals, except for nurses and midwives, in order to assess the external validity of the T-TPQ. Participants were recruited through a website in February 2020, and an anonymous online survey was conducted. Data on participants' characteristics, such as occupation type, years of experience, and gender, were collected. Finally, data from Test 1 and Test 2 were combined into a single dataset.

To examine intraclass correlation coefficient, we used the test–retest method. Medical professionals (who did not overlap with Tests 1 and 2 participants) were recruited via a website on voluntary basis, and they completed the translated T-TPQ online. Two weeks later, participants were asked to complete the translated T-TPQ again. Respondents of this test did not overlap with those who participated in Tests 1 and 2.

3.2.3 | Data analysis

Participants' characteristics are expressed as numbers and percentages, or as medians and interquartile ranges (IQR) for non-normally distributed data or means and standard deviations (*SD*) for normally distributed data. The mean score for the teamwork dimension and individual item responses are expressed as means and *SD*. Items with >50% missing responses in each dimension were excluded from the dataset. We did not perform exploratory factor analysis (EFA) because as far as we knew, T-TPQ was developed using EFA (Battles, n.d.). Our purpose was to confirm structural equivalence between two T-TPQs and assess construct validity compared with the original measurement. Therefore, we chose to perform confirmatory factor analysis (CFA) between the translated T-TPQ's structure and the previous validation study which used the original language T-TPQ items (Keebler et al., 2014). Multiple fit indices including, root-mean-square error of approximation (RMSEA), Turker–Lewis index (TLI) and comparative fit index (CFI) were used to measure the overall data-model fit. An RMSEA value <0.08 was considered acceptable. TLI and CFI values >0.95 were considered acceptable (Hu & Bentler, 1999). The adequate sample size for CFA is still being debated, and according to one study, a minimum sample size of 200 is required to obtain adequate statistical power (Hoe, 2008), whereas another study suggests that a sample size >300 is optimal for CFA (Tabachnick & Fidell, 2013). Recent studies seem to agree that a sample size of 50 is very poor, 100 is poor, 200 is average, 300 is good and 500 is very good for CFA. (DeVellis, 2016) Williams, Onsman, & Brown, 2012). Thus, we set the sample size to >500, assuming a response rate of 60% in Test 1, which recruited 800 participants. Additionally, we assessed the ceiling and floor effects corresponding to the percentage of respondents who obtained minimum (0) or maximum (5) for each question. Above 15% of respondents either minimum or maximum indicated there was a problem with validity (McHorney & Tarlov, 1995).

We used the two-way random-effects model of intraclass correlation coefficient (ICC) as an indicator of test–retest reliability. ICCs < 0.5 indicate poor reliability; ICCs between 0.5 and 0.75 indicate moderate reliability, ICCs between 0.75 and 0.9 indicate good reliability and ICCs > 0.9 indicate excellent reliability (Koo & Li, 2016). We also calculated the sample size for the intraclass correlation coefficient using test–retest method based on the ICC statistic. According to the formula recommended by Zou (2012), we set the null hypothesis at $p = .4$ (fair reproducibility); 41 observations were required to reach a significance level of .05 and test power of .08. Assuming a response rate of 65%, we selected 63 participants.

To verify the internal consistency of the dimensions of the questionnaire, we used Cronbach's α, which is the most widely used index for internal consistency (Cronbach, 1951). Cronbach's $\alpha > 0.7$ was considered to indicate acceptable to satisfactory internal consistency.

Missing data on the T-TPQ were imputed using mean substitution for dimension if there were more than 1 or 2 missing data in a dimension. Cases with 3 or more missing data in a dimension were excluded. Missing data within characteristics were excluded from the analysis.

3.3 | Ethical consideration

The study protocol was reviewed and approved by the ethics committee (#1902-1 and #1939-1). Participants were informed of the voluntary nature of the surveys and that returning completed questionnaires was considered the consent to participate.

4 | RESULTS

4.1 | Construct validity of the Japanese version of the T-TPQ

4.1.1 | Characteristics of the participants

In Test 1, of the 876 nurses and midwives, 507 (response rate: 57.9%) participants from 17 units or wards responded. Twenty-seven respondents were excluded due to missing responses for participant characteristics, and two were excluded because they did not complete more than 50% of the T-TPQ. Consequently, data from 478 participants were analysed.

In Test 2, 109 respondents were included. Thus, the total number of respondents was 587. Characteristics of respondents from both tests are shown in Table 1. Respondents included physicians, physical therapists and pharmacists, in addition to nurses and midwives; however, the majority of respondents were nurses.

TABLE 1 Characteristics of participants for Tests 1 and 2, $N = 587$

Variable	
Gender (female), *n* (%)	443(75.5)
Age group, *n* (%)	
20–24	61(10.4)
25–29	141(24.0)
30–34	93(15.8)
35–44	166(28.3)
45–55	104(17.7)
>56	22(3.8)
Years of experience, *n* (%)	
<1	29(4.9)
1–2	44(7.5)
3–5	110(18.7)
6–10	127(21.6)
11–20	145(24.7)
>20	132(22.5)
Profession, *n* (%)	
Registered Nurse	446(76.0)
Midwife	32(5.5)
Doctor	40(6.8)
Clinical engineer	26(4.4)
Physical therapist	14(0.5)
Pharmacist	10(1.7)
Medical social worker	6(1.0)
Laboratory technician	5(0.9)
Dietitian	3(0.5)
Occupational therapist	3(0.5)
Clinical radiologist	2(0.3)

4.1.2 | Construct validity of the Japanese version of the T-TPQ

The mean score, standard deviation (*SD*) and standardized factor loading for each item are shown in Table 2. Total variance value was 14.698 (41.994%). The indices of the CFA are presented in Table 3.

4.1.3 | Ceiling and floor effect in the Japanese version of the T-TPQ

The ceiling effect that was the percentage of respondents rated "5" varied from 2.9% to 27.0%. The floor effect corresponded to the percentage of respondents rated as "0" were from 0% to 7.7%. We found both Q 22 and Q 24 had ceiling effects and the percentage of respondents rated, as "5" were 17.0% and 15.7%, respectively.

4.2 | Reliability of the Japanese version of the T-TPQ

4.2.1 | Characteristics of the participants

In the test–retest method, we included 51 participants (46 nurses, 4 physicians and 1 pharmacist) who completed the test at both time points. Characteristics of the participants are shown in Table 4.

4.2.2 | Reliability of the Japanese version of the T-TPQ

The Cronbach's α for all dimensions are shown in Table 5. The Cronbach's alpha for the overall T-TPQ was 0.956, and the Cronbach's alphas for all dimensions were >0.8, indicating acceptable internal consistency. ICCs for all dimensions ranged from 0.838 to 0.957, and lowest value at 95% CI was 0.717, indicating at least moderate reliability (Table 6).

5 | DISCUSSION

In the present study, the Japanese version of the T-TPQ was developed by a multistep forward–back translation protocol, clarity and validity evaluations by experts, and pilot tests. This study is the first to formally translate the T-TPQ into Japanese and evaluate its psychometric properties. The findings of this study confirm the reliability and validity of the Japanese translated T-TPQ.

We consider the construct validity of the Japanese version of the T-TPQ to be acceptable. The RMSEA was 0.064, indicating a good fit to the original structure, and the RMSEA of our study is comparable to those reported in previous studies that translated the T-TPQ into other languages. In the Norwegian study

TABLE 2 Mean Scores, standard deviations and factor loading for T-TPQ items and dimensions $N = 587$

Teamwork dimensions and items	Items statistics		
	Mean	(*SD*)	Factor loading[a]
Team structure			
1. The skills of staff overlap sufficiently so that work can be shared when necessary	3.81	(0.77)	0.495
2. Staff are held accountable for their actions	3.86	(0.69)	0.640
3. Staff within my unit share information that enables timely decision-making by the direct patient care team	3.67	(0.74)	0.688
4. My unit makes efficient use of resources (e.g., staff supplies, equipment, information)	3.25	(0.93)	0.676
5. Staff understand their roles and responsibilities	3.80	(0.65)	0.695
6. My unit has clearly articulated goals	3.57	(0.83)	0.753
7. My unit operates at a high level of efficiency	3.12	(0.94)	0.773
Leadership			
8. My supervisor/manager considers staff input when making decisions about patient care	3.61	(0.88)	0.792
9. My supervisor/manager provides opportunities to discuss the unit's performance after an event	3.71	(0.89)	0.746
10. My supervisor/manager takes time to meet with staff to develop a plan for patient care	3.11	(0.99)	0.692
11. My supervisor/manager ensures that adequate resources (e.g., staff, supplies, equipment, information) are available	3.33	(0.93)	0.758
12. My supervisor/manager resolves conflicts successfully	3.14	(0.99)	0.776
13. My supervisor/manager models appropriate team behaviour	3.28	(0.98)	0.712
14. My supervisor/manager ensures that staff are aware of any situations or changes that may affect patient care	3.27	(0.90)	0.850
Situation monitoring			
15. Staff effectively anticipate each other's needs	3.50	(0.75)	0.765
16. Staff monitor each other's performance	3.54	(0.76)	0.792
17. Staff exchange relevant information as it becomes available	3.67	(0.72)	0.794
18. Staff continuously scan the environment for important information	3.54	(0.77)	0.773
19. Staff share information regarding potential complications (e.g., patient changes, bed availability)	3.83	(0.70)	0.713
20. Staff meets to re-evaluate patient care goals when aspects of the situation have changed	3.50	(0.82)	0.648
21. Staff correct each other's mistakes to ensure that procedures are followed properly	3.58	(0.75)	0.742
Mutual support			
22. Staff assist fellow staff during high workload	3.93	(0.73)	0.618
23. Staff request assistance from fellow staff when they feel overwhelmed	3.74	(0.73)	0.353
24. Staff caution each other about potentially dangerous situations	3.94	(0.68)	0.698
25. Feedback between staff is delivered in a way that promotes positive interactions and future change	3.50	(0.80)	0.752
26. Staff advocate for patients even when their opinion conflicts with that of a senior member of the unit	3.32	(0.82)	0.634
27. When staff have a concern about patient safety, they challenge others until they are sure the concern has been heard	3.53	(0.75)	0.629
28. Staff resolve their conflicts, even when the conflicts have become personal	3.34	(0.79)	0.661

(Continues)

TABLE 2 (Continued)

Teamwork dimensions and items	Items statistics		
	Mean	(*SD*)	Factor loading[a]
Communication			
29. Information regarding patient care is explained to patients and their families in lay term	3.92	(0.58)	0.552
30. Staff relay relevant information in a timely manner	3.75	(0.67)	0.711
31. When communicating with patients, staff allow enough time for questions	3.57	(0.76)	0.553
32. Staff use common terminology when communicating with each other	3.84	(0.65)	0.538
33. Staff verbally verify information that they receive from one another	3.44	(0.80)	0.631
34. Staff follow a standardized method of sharing information when handing off patients	3.51	(0.81)	0.645
35. Staff seek information from all available sources	3.73	(0.71)	0.706

[a]Factor loading are standardized.

TABLE 3 CFA fit indices

	Indices
χ^2	1,782.633, $p < .001$
RMSEA	0.062, 90%CI [0.059–0.065]
CFI	0.895
TLI	0.886

Abbreviations: CFA, confirmatory factor analysis; CFI, comparative fit index; CI, confidential interval; RMSEA, root square error of approximation; TLI, Turker–Lewis index.

TABLE 4 Characteristics of participants for test–retest method ($N = 51$)

Gender (female), *n* (%)	25 (49.0)
Age group, *n* (%)	
20–24	1 (2.0)
25–29	3 (5.9)
30–34	16 (31.4)
35–44	21 (41.2)
45–55	10 (19.6)
>56	0 (0)
Years of experience, *n* (%)	
<1	0 (0)
1–2	1 (2.0)
3–5	2 (3.9)
6–10	17 (33.3)
11–20	18 (35.3)
>20	13 (25.5)
Profession, *n* (%)	
Registered Nurse	46 (90.2)
Physician	4 (7.8)
Pharmacist	1 (2.0)

(Ballangrud et al., 2017), the RMSEA was 0.056, which was slightly lesser than that reported in this study. However, the Korean study (Hwang & Ahn, 2015) reported a similar RMSEA (0.067). These results indicate that the original version of the T-TPQ has a rigorous structure, and the Japanese version of T-TPQ maintained its validity. Although TLI and CFI in this study were below the recommended values, they were closer to the recommended values compared to the Norwegian study (TLI 0.878, CFI 0.833) (Ballangrud et al., 2017). About these indexes, a few factors need to be considered when interpreting the results of the analysis. The fit indices may be affected by differences in the kind of population; in the previous larger study (Keebler et al., 2014), the participants were multi-professional healthcare professionals, but all of them were military personnel. Therefore, these differences might have contributed to our findings. About RMSEA, which is the most robust index of CFA (MacCallum & Austin, 2000), we believe that the construct validity of the Japanese version of T-TPQ was acceptable despite the lower values of TLI and CFI.

The ceiling effect was found in two items. As far as we know, this is the first study to examine ceiling and floor effects of T-TPQ of any language including the original measurement; thus, we could not make a decision which was problematic, the original T-TPQ, translation, or higher safety culture in this population. Further research is required to investigate the ceiling and floor effects of the original T-TPQ.

The ICC of the translated T-TPQ as an index of test–retest reliability was at least moderate; however, all estimates were >0.8. Thus, the Japanese version of the T-TPQ has good test–retest reliability. Test–retest reliability of the instrument was also assessed in the Norwegian study (Ballangrud et al., 2017), and they reported ICCs for the dimensions ranging from 0.672 to 0.852. The differences may be due to the sample size (total sample of 244), which was more than our study. However, the internal consistency of the T-TPQ

TABLE 5 Cronbach's alpha in each dimensions

T-TPQ subscales	Cronbach's alpha	95% CI	*F* test value	*p* value
Team structure	0.851	0.717–0.908	6.1310	<.001
Leadership	0.917	0.924–0.975	23.019	<.001
Situation monitoring	0.897	0.873–0.959	13.773	<.001
Mutual support	0.860	0.856–0.953	12.196	<.001
Communication	0.817	0.764–0.923	7.413	<.001

Abbreviation: CI, confidential interval.

TABLE 6 ICCs for test–retest method for five dimensions

T-TPQ subscales	ICC	95% CI	*F* test value	*p* value
Team structure	0.838	0.717–0.908	6.1310	<.001
Leadership	0.957	0.924–0.975	23.019	<.001
Situation monitoring	0.927	0.873–0.959	13.773	<.001
Mutual support	0.918	0.856–0.953	12.196	<.001
Communication	0.865	0.764–0.923	7.413	<.001

Abbreviations: CI, confidential interval; ICC, intraclass correlation coefficient.

is acceptable, because the lowest Cronbach's alpha at 95% CI was 0.717, which is considered acceptable (Cronbach, 1951).

This study has some advantages. First, we used a multistep back translation method to ensure that the items are clearly understood by the healthcare professionals. Especially, several medical professions participated in the translation process that improved the quality of translation. Second, the sample size for Test 1 was relatively higher than those in previous translation–validation studies of the T-TPQ, indicating a more conclusive finding from the present study. Additionally, in Test 1, we obtained responses from nurses and midwives, although participants were recruited from a single university hospital. Third, we also collected data from multi-professionals in Test 2, and we combined data from Tests 1 and 2, so that our findings represent healthcare professionals from various settings in a hospital. Our findings suggest that the Japanese version of the T-TPQ is useful for not only nursing staff but also other medical professionals. This may enable researchers to evaluate teamwork in workplace and contribute to improvement.

The present study also has some limitations. First, we made efforts to follow the guidelines published by Sousa (Sousa & Rojjanasrirat, 2011); however, we did not completely follow the guidelines due to insufficient human resources. These include the questionnaire being back translated by one person (two people recommended) and forward translated by two persons (one of the translators did not have adequate knowledge of medical terminology). These limitations may have influenced the adequacy of our translation. To avoid losing equivalence of meaning owing to possible differences in hospital culture, some items were translated by a bilingual medical professional who worked in both Japan and the United States. However, we believe that despite these limitations, the findings of our study contribute to enhancing performance and patient safety in hospital. Further, we believe that the Japanese version of the T-TPQ ensured adequate equivalence of meaning with the original T-TPQ. Further research is needed to investigate the ceiling and floor effects of the original T-TPQ and revise the entire T-TPQ using EFA. Considering this analysis, T-TPQ will be a more specific and accurate measurement for perceived teamwork.

6 | CONCLUSION

Our findings suggest that the Japanese version of the T-TPQ is acceptable in terms of validity and reliability. T-TPQ can be compared internationally amongst countries. Teamwork can be held at the hospital and unit levels. This scale may help assess teamwork in hospital settings, which may facilitate the improvement of care quality.

ACKNOWLEDGEMENTS

The authors acknowledge the members of the expert panel and participants for their support and cooperation.

CONFLICT OF INTEREST

No conflict of interest has been declared by the authors.

AUTHOR CONTRIBUTIONS

All authors listed meet the authorship criteria according to the latest guidelines of the International Committee of Medical Journal Editors are in agreement with the manuscript.

ORCID

Takeshi Unoki https://orcid.org/0000-0002-7778-4355
Yujiro Matsuishi https://orcid.org/0000-0001-6388-9263

REFERENCES

Ballangrud, R., Husebø, S. E., & Hall-Lord, M. L. (2017). Cross-cultural validation and psychometric testing of the Norwegian version of the TeamSTEPPS® teamwork perceptions questionnaire. *BMC Health Services Research*, *17*(1), 1–10. https://doi.org/10.1186/s12913-017-2733-y

Battles, J.. *TeamSTEPPS teamwork perception questionnaire (T-TPQ) manual*. Retrieved from https://www.ahrq.gov/teamstepps/instructor/reference/teamperceptionsmanual.html

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–333. https://doi.org/10.1007/BF02310555

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage.

Fujita, S., Seto, K., Ito, S., Wu, Y., Huang, C. C., & Hasegawa, T. (2013). The characteristics of patient safety culture in Japan, Taiwan and the United States. *BMC Health Services Research*, *13*. Article no: 20. https://doi.org/10.1186/1472-6963-13-20

Hoe, S. L. (2008). Issues and procedures in adopting Structural Equation Modeling Technique. *Journal of Applied Quantitative Methods*, *3*, 76–83.

Hu, L.-T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hwang, J. I., & Ahn, J. (2015). Teamwork and clinical error reporting among nurses in Korean hospitals. *Asian Nursing Research*, *9*(1), 14–20. https://doi.org/10.1016/j.anr.2014.09.002

Ito, S., Seto, K., Kigawa, M., Fujita, S., Hasagawa, T., & Hasegawa, T. (2011). Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. *BMC Health Services Research*, *11*. Article no: 28. https://doi.org/10.1186/1472-6963-11-28

Keebler, J. R., Dietz, A. S., Lazzara, E. H., Benishek, L. E., Almeida, S. A., Toor, P. A., … Salas, E. (2014). Validation of a teamwork perceptions measure to increase patient safety. *BMJ Quality and Safety*, *23*(9), 718–726. https://doi.org/10.1136/bmjqs-2013-001942

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*(3), 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: Are available health status surveys adequate? *Quality of Life Research*, *4*(4), 293–307. https://doi.org/10.1007/BF01593882

Parker, A. L., Forsythe, L. L., & Kohlmorgen, I. K. (2019). TeamSTEPPS®: An evidence-based approach to reduce clinical errors threatening safety in outpatient settings: An integrative review. *Journal of Healthcare Risk Management*, *38*(4), 19–31. https://doi.org/10.1002/jhrm.21352

Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268–274. https://doi.org/10.1111/j.1365-2753.2010.01434.x

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). New York: Pearson.

Unoki, T., Hamamoto, M., Sakuramoto, H., Shirasaka, M., Moriyasu, M., Zeng, H., & Fujitani, S. (2020). Influence of mutual support and a culture of blame among staff in acute care units on the frequency of physical restraint use in patients undergoing mechanical ventilation. *Acute Medicine & Surgery*, *7*(1), 1–6. https://doi.org/10.1002/ams2.479

Weller, J., Boyd, M., & Cumin, D. (2014). Teams, tribes and patient safety: Overcoming barriers to effective teamwork in healthcare. *Postgraduate Medical Journal*, *90*(1061), 149–154. https://doi.org/10.1136/postgradmedj-2012-131168

Williams, B., Onsman, A., & Brown, T. (2012). EDUCATION exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3), 1–13. Retrieved from http://ro.ecu.edu.au/jephc/vol8/iss3/1

Zou, G. Y. (2012). Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Statistics in Medicine*, *31*(29), 3972–3981. https://doi.org/10.1002/sim.5466

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Unoki T, Matsuishi Y, Tsujimoto T, et al. Translation, reliability and validity of Japanese version the TeamSTEPPS® teamwork perceptions questionnaire. *Nursing Open*. 2021;8:115–122. https://doi.org/10.1002/nop2.609