

<Original Article>

Development of a Questionnaire to Assess Bangladeshi Government Primary School Teachers' Training Needs for Inclusive Education

Md Abu Bakor Siddik* · Norimune Kawai**

Abstract

Bangladesh initiated inclusive education in state schools in 2010 and introduced reforms to the teacher education program in 2015 for its effective implementation. However, studies have found that the current teacher training is insufficient for government primary school teachers in Bangladesh. It is important to understand their needs to design more effective teacher training programs. This study developed a questionnaire to reveal Bangladeshi government primary school teachers' training needs for inclusive education. Development of the questionnaire was conducted in three stages: For Phase 1, the domains of the teachers' needs were identified; in Phase 2, domain items were developed and tested; and in Phase 3, the psychometric properties of the tool were examined for reliability and validity. This research mainly focuses on the results of the reliability and validity testing of the questionnaire. The questionnaire originally comprised 48 items within nine subscales. Eventually, 37 items were selected within eight subscales after validity and reliability testing demonstrated good reliability, internal consistency, and convergent validity. All but one of the subscales are distinct and unidimensional.

Keywords: Inclusive education, Primary school, Teacher training, Bangladesh

Introduction

Although government primary schools (GPSs) in Bangladesh began to include children with special needs in 2005, inclusive education really began after implementing the National Education Policy 2010 (Ahsan, Deppeler, & Sharma, 2013). The in-service training program was also reformed in 2015 so that GPS teachers could learn to work appropriately with students with special needs (Ahsan, Tasnuba, Akter, Islam, & Miah, 2016; UNICEF, 2014). From 2007 to 2012, the Directorate of Primary Education (DPE) conducted training for GPS head-teachers, education officers, and teacher trainers in inclusive education with funding from the World Bank and the Japan International Cooperation Agency under Primary

Education Development Programs II and III (Ahsan, Sharma, & Deppeler, 2012).

However, researchers have found that GPSs face many challenges in enrolling students with special needs, school management, and breaking down social barriers, including teacher education (Ahsan et al., 2012; Malak, 2013). Studies have also found that existing prerequisite training programs are not working effectively to improve teachers' attitudes and competencies (Siddik & Kawai, 2018). Therefore, it is important to understand teacher training needs for inclusive education, so both researchers and teachers can develop new teacher education programs to incorporate inclusive education.

Inclusive education is still a new approach in many GPSs, and although it has been more than a decade since it was introduced, there have been insufficient training opportunities (Ahsan et al., 2016). In Bangladesh, a GPS teacher candidate does not need to have a degree in education to teach. Once employed by a GPS, teachers receive in-service training and work toward their Certificate in

* Educational Development Course, Graduate School for International Development and Cooperation, Hiroshima University, Japan.

** Division of Educational Sciences, Graduate School of Humanities and Social Sciences, Hiroshima University, Japan.

Education or Diploma in Primary Education. In addition to this training, GPS teachers take part in short, subject-based, and refresher courses. However, as inclusive education is still relatively unknown among GPS teachers, it is challenging to understand their training needs when they do not know what they are. Consequently, the current study found a challenge to collect data from GPS teachers by interviewing them, as they could not provide enough information. Moreover, there are no globally recognized tools to measure GPS teachers' training needs, and a valid and reliable questionnaire instrument is needed.

The purpose of this study was to conduct an initial psychometric evaluation of a questionnaire to assess Bangladeshi GPS teachers' training needs for inclusive education. The psychometric evaluation aimed to examine variance and nonresponse rates for the questionnaire's items, each subscale's unidimensionality, and construct validity.

Methods

Development of the instrument was conducted in three stages. In Phase 1, 18 domains of GPS teachers' training needs were identified from a literature review. In Phase 2, domain items were developed and tested for clarity and relevance to GPS teachers' training needs for inclusive education. In Phase 3, the psychometric properties of the tool were examined with a self-administered pilot study.

In Bangladesh, 15 types of educational institutions provide primary education (DPE, 2014). The current study is focused on only one kind of institution called a GPS. According to Bangladesh's National Education Policy, 2010, all primary schools are supposed to accept children with special needs. GPS teachers in Bangladesh comprised the population for this study. For the pilot study, respondents were selected from three divisions in Bangladesh.

Phase 1: Identification of Domains of Government Primary School Teachers' Training Needs for Inclusive Education

The first phase in developing this questionnaire was to outline its purpose and indicate GPS teachers'

training needs for inclusive education (Streiner, Norman, & Cairney, 2015). Some questionnaires measure a single phenomenon, while others measure multiple phenomena such as capacities (Polit & Beck, 2006). The current questionnaire has multiple purposes for identifying GPS teachers' training needs according to teacher training components (Younas & Porr, 2018). To recognize the phenomenon, literature reviews, theories, and concept analyses were used to develop a survey questionnaire (Younas & Porr, 2018). Eighteen domains were identified in the literature review. They were sorted according to professional knowledge, professional practice, and professional engagement, and the contents sorted from secondary data were categorized into teacher education components (Ahsan et al., 2012; Dyson, Plunkett, & McCluskey, 2018).

Phase 2: Development of the Items

From the 18 identified domains of teachers' training needs, nine objectives were developed. These have been listed in Columns 1 and 2 of Table 1.

Phase 3: Pilot Study and Psychometric Testing

Psychometric testing establishes the validity and reliability of a questionnaire (Morgado, Meireles, Neves, Amaral, & Ferreira, 2017). This pilot study measured the content validity, construct validity, and internal consistency as part of the reliability validity. The face validity of the questionnaire was also assessed. Common validity and reliability measures were used.

Data Analysis

Descriptive statistics were used to assess the demographics and characteristics of the study participants. The Aiken V value model was run for measuring content validity using Microsoft Excel 2013 (Aiken, 1985). Moreover, SPSS Statistics 25.0 (IBM, 2017) was used to measure the correlation between items, subscales, and total values. Also, principal component analysis (PCA) was run to measure the commonalities of individual items.

Bangladeshi Government Primary School Teachers' Training Needs

Table 1
Item-Total Score Correlations and Communalities for Pilot and Revised Scales

Objectives (1)	Contents for training on inclusive education (2)	Questions (3)	Comments (4)	Correlation with a total score (5)	Correlation with subscale total (6)	Communalities (7)	Remarks (8)
To measure government primary school (GPS) teachers' attitudes towards inclusive education.	Motivation for teaching children with special needs	1. Students who have difficulty expressing their thoughts verbally should be in regular classes.	No change.	0.299 [†]	0.580 [†]	0.758	Accepted
		2. Students who are less attentive should be in regular classes.	No change.	0.499 [†]	0.747 [†]	0.895	Accepted
		3. Students who require communicative technologies (e.g., Braille/sign language) should be in regular classes.	No change.	0.559 [†]	0.881 [†]	0.886	Accepted
		4. Students who need an individualized academic program (IEP) should be in regular classes.	No change.	0.525 [†]	0.831 [†]	0.891	Accepted
To measure GPS teachers' sentiments towards inclusive education.	Positive attitude towards children with special needs	5. I dread the thought that if I stay with a person with disabilities, I could eventually end up with a disability.	The statement was not clear, according to the expert panel's review. Therefore, this statement has been elaborated to make it clear.	0.226	0.618 [†]	0.834	Rejected
		6. I tend to keep contact with people with disabilities, and I finish this contact as quickly as possible.	No change.	0.223	0.417 [†]	0.761	Rejected
		7. I would feel terrible if I had a disability.	No change.	0.385 [†]	0.634 [†]	0.838	Accepted
		8. I am afraid to look straight in the face of a person with a disability.	No change.	0.553 [†]	0.668 [†]	0.798	Accepted
		9. I find it difficult to overcome my initial shock when meeting people with severe physical disabilities.	No change.	0.354 [†]	0.708 [†]	0.841	Accepted
To measure GPS teachers' efficacy in using inclusive instructions in the classroom for children with special needs.	Classroom management for children with special needs Teaching in a large size class	10. I can provide individual attention to students with special needs without hampering my class.	No change.	0.640 [†]	0.781 [†]	0.875	Accepted
		11. I am well aware of seating arrangements in an inclusive classroom.	No change.	0.619 [†]	0.687 [†]	0.838	Accepted
		12. I am confident in my ability to get students to work together in pairs or small groups.	No change.	0.596 [†]	0.831 [†]	0.878	Accepted
		13. I can manage tasks in time for inclusive classes.	No change.	0.603 [†]	0.732 [†]	0.677	Rejected
		14. I am confident in my ability to prevent disruptive behavior in the classroom.	No change.	0.638 [†]	0.740 [†]	0.884	Accepted
		15. I am able to get children to follow classroom rules.	No change.	0.715 [†]	0.884 [†]	0.927	Accepted
		16. I am confident when dealing with students who are physically aggressive.	No change.	0.734 [†]	0.831 [†]	0.830	Accepted
		17. I can teach in a large size class.	No change.	0.671 [†]	0.790 [†]	0.832	Accepted
		18. I can use group teaching and peer teaching to teach my students.	No change.	0.425 [†]	0.517 [†]	0.926	Accepted
Assessment and evaluation	Assessment of children with special needs	19. I am aware of using formative assessment in an inclusive class.	No change.	0.464 [†]	1.000 [†]	0.864	Accepted
		20. I am aware of using summative assessment in an inclusive class.	No change.	0.580 [†]	0.734 [†]	0.858	Accepted
		21. I believe assessment strategies may vary among students based on their special needs.	No change.	-0.017	-0.111	0.826	Rejected
		22. I can use various assessment strategies (e.g., portfolio assessment, modified tests, performance-based assessments).	No change.	0.575 [†]	0.359 [†]	0.816	Accepted
To measure GPS teachers' efficacy in collaborating with stakeholders to establish inclusive education.	Collaboration inside the school for inclusive education	23. I am able to work jointly with other professionals and staff (e.g., aides, other teachers) to teach students with disabilities in the classroom.	No change.	0.588 [†]	0.849 [†]	0.871	Accepted
		24. I can collaborate with other professionals (e.g., itinerant teachers or speech pathologists) in designing educational plans for students with disabilities.	No change.	0.599 [†]	0.822 [†]	0.823	Accepted
	Collaboration outside the school for inclusive education	25. I can assist families in helping their children do well in school.	No change.	0.559 [†]	0.890 [†]	0.824	Accepted
		26. I am confident in informing others who know little about laws and policies relating to the inclusion of students with disabilities.	No change.	0.653 [†]	0.822 [†]	0.851	Accepted

Objectives (1)	Contents for training on inclusive education (2)	Questions (3)	Comments (4)	Correlation with a total score (5)	Correlation with subscale total (6)	Communalities (7)	Remarks (8)	
To measure GPS teachers' concerns about resources for inclusive education.	Manage local resources for inclusive education	27. My school will not have enough funds to implement integration successfully.	No change.	0.152	0.821 [~]	0.805	Rejected	
		28. An inadequate level of paraprofessional staff will be available to support integrated students (e.g. speech therapist, physiotherapist, occupational therapist).	No change.	0.099	0.700 [~]	0.853	Rejected	
		29. My school will have difficulty accommodating students with various types of disabilities due to inappropriate infrastructure, e.g., architectural barriers.	No change.	-0.039	0.700 [~]	0.919	Rejected	
		30. Inadequate resources or special teachers will be available to support integration.	No change.	0.130	0.703 [~]	0.674	Rejected	
		31. My school will not have adequate special education instructional materials and teaching aids, e.g., Braille.	No change.	0.216	0.810 [~]	0.777	Rejected	
		32. Inadequate administrative support will be available to implement the integration program.	No change.	0.136	0.757 [~]	0.744	Rejected	
		33. My school management committee positively considers my opinions.	No change.	0.425 [~]	0.089	0.912	Rejected	
To measure GPS teachers' efficacy in teaching approaches for children with special needs.	Teaching approaches for children with special needs	34. I can make my expectations clear about student behavior.	No change.	0.643 [~]	0.781 [~]	0.818	Accepted	
		35. I can provide an alternative explanation, e.g. when students are confused.	No change.	0.663 [~]	0.712 [~]	0.872	Accepted	
		36. I am confident in designing learning materials and tasks to meet the individual needs of students with disabilities.	No change.	0.659 [~]	0.801 [~]	0.891	Accepted	
		37. I can accurately measure students' levels of comprehension of the content of what I have taught.	No change.	0.356 [~]	0.598 [~]	0.823	Accepted	
To measure GPS teachers' efficacy in dealing with disruptive behaviors when establishing inclusive education.	Teaching aids for children with special needs	38. I can use teaching materials (e.g. multisensory teaching aids) that are important for children with special needs.	No change.	0.712 [~]	0.642 [~]	0.835	Accepted	
	Sign language Braille	39. I think it is essential for GPS teachers to use sign language in an inclusive classroom.	No change.	0.338 [~]	0.744 [~]	0.818	Accepted	
		40. I think it is essential for GPS teachers to use Braille in an inclusive classroom.	No change.	0.422 [~]	0.764 [~]	0.718	Accepted	
To measure GPS knowledge about inclusive education.	Education policy and Acts for children with special needs	41. I know the national laws and policies related to the education of children with special needs.	No change.	0.564 [~]	0.867 [~]	0.920	Accepted	
		42. I know the international laws and policies related to the education of children with special needs.	No change.	0.563 [~]	0.845 [~]	0.863	Accepted	
	Definition, objectives, and purpose of inclusive education from the national perspective	43.	I know the definitions and objectives of inclusive education from the Bangladeshi perspective.	No change.	0.385 [~]	0.730 [~]	0.820	Accepted
			I know the definitions and objectives of inclusive education from an international perspective.	No change.	0.560 [~]	0.853 [~]	0.884	Accepted
	Introduction to children with special needs	45.	I know the characteristics of special needs children.	No change.	0.390 [~]	0.737 [~]	0.901	Accepted
			I know about diversity.	No change.	0.295 [~]	0.568 [~]	0.838	Accepted
	Understanding diversity and individual needs	47.	I know about students' individual learning needs.	No change.	0.517 [~]	0.663 [~]	0.869	Accepted
			I know how I can make my school accessible to children with special needs.	No change.	0.618 [~]	0.701 [~]	0.909	Accepted

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.

Domains of Government Primary School Teachers' Training Needs for Inclusive Education

The literature review located 97 papers related to teacher training for inclusive education in Bangladesh, of which 25 papers noted teacher training needs. From these 25 papers, 18 recognized GPS teachers' training needs. The identified training needs are as follows: motivation for teaching children with special needs; positive attitude towards children with special needs; classroom management for children with special needs; teaching in a large classroom; assessment of children with special needs; collaboration inside the school for inclusive education; collaboration outside the school for inclusive education; management of local resources for inclusive education; teaching approaches for children with special needs; teaching aids for children with special needs; sign language; Braille; policy and Acts related to education for children with special needs; definition, objectives, and purpose of inclusive education from the national perspective; definition, objectives, and purpose of inclusive education from an international perspective; introduction to children with special needs; understanding diversity and individual needs; and making the school accessible for all. These needs are presented in Table 1, Column 2.

Development of Items

From Phase 1 findings, 48 items were selected for the proposed questionnaire. Four items (1–4) were adapted from the attitude subscale of the Sentiments, Attitudes, and Concerns about Inclusive Education Revised (SACIE-R) scales to measure the impact on GPS teachers of inclusive education (Forlin, Earle, Loreman, & Sharma, 2011). Five items (5–9) were adopted from the sentiment subscale of the SACIE-R scales to measure GPS teacher sentiments towards inclusive education (Forlin et al., 2011). Four items (12, 14, 15, and 16) were adapted from the inclusive classroom management part of the Teacher Efficacy for Inclusive Practice (TEIP) scale (Sharma, Loreman, & Forlin, 2012). Moreover, five items (10, 11, 13, 17, and 18) were created to measure GPS teachers' efficacy in using inclusive instructions in the classroom for children with special needs. One item (22) was adopted from the

TEIP scale to measure teacher training needs in relation to assessment and evaluation, along with three additional items (19, 20, and 21) generated from the literature review. Four items (23, 24, 25, and 26) were adapted from the TEIP scale to measure GPS teachers' efficacy in collaboration with stakeholders to establish inclusive education (Sharma et al., 2012). Seven items (27, 28, 29, 30, 31, 32, and 33) were adapted from the Concerns about Integrated Education (CIE) scale to measure GPS teachers' concerns about resources for inclusive education (Sharma & Desai, 2002). Four items (34, 35, 36, and 37) were collected from the TEIP scale to measure the efficacy of GPS teachers' teaching approaches for children with special needs (Sharma et al., 2012). Three items (38, 39, and 40) were created from the literature review findings to measure the efficacy of GPS teachers' dealing with disruptive behaviors when establishing inclusive education. Moreover, eight items (41, 42, 43, 44, 45, 46, 47, and 48) were created to measure teachers' knowledge about inclusive education, with these items arising from the literature review findings. The objectives of all items and the subscales' objectives are presented in Table 1, Columns 1, 2, and 3.

Pilot Study and Psychometric Testing

The main questionnaire was written in English, translated into Bengali, and checked by a university professor working in the education field for more than 15 years. The translated questionnaire was back-translated into English by two experts in the education sector. Eventually, the Bengali translation was evaluated against the English questionnaire. A pilot study was held after the finalization of the translation. A Google form was developed to collect data from GPS teachers. The form's link was disseminated among GPS teachers by the Primary Teachers Training Institute (PTI) instructors in three divisions. In the pilot study, 46 responses responded.

Face validity

Face validity is the only qualitative measure of a questionnaire's validity. This includes expert opinions and reviews of the scales and items. After experts review the questionnaire, the results are

compared, and the questionnaire is modified. The main limitation of face validity is that it is highly subjective (Bolarinwa, 2015; Younas & Porr, 2018).

Five experts in the education sector were asked to measure the face validity of the proposed questionnaire, considering the clarity of the wording, the likelihood of the target audience being able to answer the questions, the layout, and the style (Parsian & Dunning, 2009). Four reviewers were able to return their reviews on time. Among the reviewers, one was a university teacher in the education faculty and had worked in education for more than 15 years. Two reviewers were working for the PTI as instructors. The fourth was a researcher who had worked in the education sector for more than 12 years.

The expert panel suggested changes to some wording and removing two items on the experience of teaching children with special needs, which were on the face sheet. In Table 1, Column 4 shows the expert s' remarks.

Content validity

For content validity, multiple experts evaluate the scale, its items, the content (which is the conceptual definition of the phenomenon), its attributes, and the items selected for its operationalization. Unlike face validity, content validity is more structured, and validity can be determined through expert consensus (Bolarinwa, 2015; Younas & Porr, 2018).

An expert panel of seven members measured the content validity of the proposed questionnaire. Among the members, three were education officers, and four were PTI instructors.

In this study, the following formula, which was proposed by Aiken (1985), was used to measure content validity:

$$\bullet V = \frac{\sum s}{[n(c-1)]}$$

$$\bullet s = r - lo$$

Where,

- lo = lowest validity rating (1)
- c = highest validity rating (5)
- r = number given by the experts

According to Aiken (1985), right-tail probabilities (p) for selected values of the validity coefficient (V)

set an error rate for the questionnaire for the current study at 5% ($p < 0.05$). Moreover, from Aiken's table standard of V coefficient, the value of V for seven (7) experts with five (5) rating scales is minimal 0.75 ($V \geq 0.75$). Within that measurement, all of the questions and statements in the proposed questionnaire were more than 0.75, except one statement in the demographic information section related to teachers' disabilities. Therefore, that question was omitted from further study.

Construct validity

Construct validity comprises discriminant/convergent validity and factorial validity and measures a scale's consistency, conceptualization, and the underlying theory by compiling existing research on the subject (Kimberlin & Winterstein, 2008). In discriminant/convergent validity, a scale is used to compare the instrument with another instrument derived from the same theory, and their correlation is then determined. If the scale and instrument are related each other, they are convergently valid; however, if they are not, they are divergently valid (Streiner et al., 2015).

Construct validity is directly concerned with one variable's theoretical relationship to other variables (DeVellis, 1991). There are many ways to measure construct validity with an analysis of variance (ANOVA), correlation coefficients, and factor analysis (Roberts & Priest, 2006). Factor analysis uses several factors to determine if the characteristics relate to each other (Bryman & Cramer, 2004).

Factor analysis is especially useful for exploring the relationships among large numbers of variables, extricating them, and identifying clusters of variables that are intricately linked together (Roberts & Priest, 2006). However, in the current pilot study, there were a total of 48 questions on the Likert scale, and the number of respondents was only 46. In this case, the number of participants was fewer than the number of items, and factor analysis was not an appropriate method to measure construct validity for the current pilot study (Field, 2009). Therefore, the correlation coefficients among the subscales and the correlation between the total and single item values were calculated to measure

construct validity (Bolarinwa, 2015; Kudsi, Fenlon, Hill, & Baysan, 2020; Sangoseni, Hellman, & Hill, 2013; Torkian, Shahesmaeili, Malekmohammadi, & Khosravi, 2020; Younas & Porr, 2018).

Results

In the current study, a total of nine subscales were used to collect data from GPS teachers. Each subscale has a different objective to measure GPS teachers' training needs for inclusive education. Correlations between the total score and every single item were analyzed to measure the current questionnaire's construct validity. In the analysis, items number 5, 6, 21, 27, 28, 29, 30, 31, 32 were not significantly correlated to the questionnaire's total score. According to the analysis results, those items did not directly relate to the total score. However, within the subscale, the correlations are statistically robust, except for 21 and 33. Columns 5 and 6 in Table 1 show the results.

Item 21 was: *I believe assessment strategies may vary among students based on their special needs.* Furthermore, item 33 was: *My school management committee positively considers my opinions.* Item 21 belonged to the assessment and evaluation for children with special needs subscale, while item 33 belonged to the concern about resources for the inclusive education section. Based on these analyses, items 21 and 33 were deleted.

Moreover, correlations between subscales were also measured to see the relationships between the subscales (Younas & Porr, 2018). In Table 1, the relationships among the subscales are shown. The subscale F (*Teachers' concerns about resources for inclusive education*) was not correlated with the other subscales because these subscale questions were much more related to school and not causally related to teachers' knowledge or skills. This subscale was divergently valid with the other scales. Subscale A (*Teachers' attitudes towards inclusive education*) was divergently valid with subscales B and D, and convergently valid with subscales C, E, G, H, and I. Subscale B (*Teachers' sentiments towards inclusive education*) was divergently valid with subscales D and I, and convergently valid with C, E, G, and H subscales. Subscale C (*Teachers' efficacy in*

using inclusive instructions in the classroom for children with special needs) was convergently valid with D, E, G, and H subscales. Subscale D (*Assessment and evaluation*) was divergently valid with subscale H and convergently valid with E, G, and I subscales. Subscale E (*Teachers' efficacy in collaborating with stakeholders to establish inclusive education*) was divergently valid with subscale I and convergently valid with G and H subscales. Subscale G (*Teachers' efficacy in teaching approaches for children with special needs*) was convergently valid with H and I subscales. Subscale H (*Teachers' efficacy in dealing with disruptive behaviors for establishing inclusive education*) was convergently valid with I (*Knowledge for inclusive education*) subscale.

According to the results described above, the F subscale has the most divergent value with the other eight subscales among the nine subscales. Moreover, this scale item was also divergently valid (-0.039 to 0.216) with the questionnaire's total score. Therefore, this subscale was removed from the proposed questionnaire.

Reliability measured by internal consistency

Internal consistency is the most commonly used measure of reliability. A researcher determines internal consistency by collecting data and then analyzing them using interitem correlation, the Kuder-Richardson index, or Cronbach's alpha (Bolarinwa, 2015).

Cronbach's alpha is widely used in educational research to measure the reliability of instruments, especially for the data collected by a Likert scale (Oluwatayo, 2012). Moreover, Cronbach's alpha takes into consideration the variance of each item (Cronbach, 1951). Cronbach's alpha coefficient was used to evaluate the reliability of the measurements for the current pilot study. The results show that the reliability of the total scale measurements was remarkably high (0.913). The reliability of 0.913 means that 91.3 percent of the observed score variability was valid, and 8.7 percent was due to error (Field, 2009; Roberts & Priest, 2006). Moreover, all the nine subscales' reliability measurements are from 0.492-0.9, which means the subscales' reliability was also acceptable. Table 2 shows the Cronbach's alpha measurement values.

Table 2
Correlations among the subscales, including Cronbach's alpha value

Correlations	A	B	C	D	E	F	G	H	I	Total
Item number in subscales	4	5	9	4	4	7	4	3	8	48
A. Teachers' attitudes towards inclusive education.	-	0.199	0.517**	0.21	0.588**	0.223	0.358*	0.382**	0.434**	
B. Teachers' sentiments towards inclusive education.	0.199	-	0.401**	0.196	0.374*	-0.006	0.514**	0.387**	0.264	
C. Teachers' efficacy in using inclusive instructions in a classroom for children with special needs.	0.517**	0.401**	-	0.445**	0.620**	-0.076	0.645**	0.454**	0.481**	
D. Assessment and evaluation	0.21	0.196	0.445**	-	0.371*	-0.039	0.338*	0.223	0.300*	
E. Teachers' efficacy in collaborating with stakeholders to establish inclusive education.	0.588**	0.374*	0.620**	0.371*	-	-0.064	0.643**	0.505**	0.264	
F. Teachers' concerns about resources for inclusive education.	-0.223	-0.006	-0.076	-0.039	-0.064	-	0.254	0.367*	-0.119	
G. Teachers' efficacy in teaching approaches for children with special needs.	0.358*	0.514**	0.645**	0.338*	0.643**	0.254	-	0.555**	0.347*	
H. Teachers' efficacy in dealing with disruptive behaviors for establishing inclusive education.	0.382**	0.387**	0.454**	0.223	0.505**	0.367*	0.555**	-	0.375*	
I. Teachers' knowledge about inclusive education.	0.434**	0.264	0.481**	0.300*	0.264	-0.119	0.347*	0.375*	-	
Cronbach's alpha	0.766	0.579	0.9	0.662	0.856	0.8	0.684	0.492	0.886	0.913

Note: Significant correlations are presented in bold characters.

Commonalities show how common the respondents' answers are to the questions (Field, 2009). According to Field (2009), with 30 or more variables and communalities greater than 0.7 for all variables, different solutions are unlikely. However, with fewer than 20 variables and any low communalities (< 0.4), differences can occur. In the current pilot study, among the 48 items, only two items have communalities of less than 0.7. Those two items were 13 (*I can manage tasks in time for inclusive classes*) and 30 (*Resources and teacher numbers are for supporting inclusive education*).

According to the result of the analysis, both items (13 and 30) were eliminated. Column 7 in Table 1 shows the communalities of the 48 items used for the current study.

In order to increase the validity and reliability of the questionnaire, 11 items were deleted. Items 5 and 6 were deleted due to low correlations with the total score, even though those items are convergently valid as a subscale. Moreover, item 21 was divergently valid with the total scale and subscale too. Therefore, those items were removed from the questionnaire. Item 13 had a low

commonality that showed it was not clear to the respondent; therefore, it was deleted.

Items 27, 28, 29, 30, 31, and 32 belong to the *Teachers' concern about resources for inclusive education* subscale. These six items are divergently valid with the total scale and subscales. Moreover, another item (33) in this subscale was convergently valid with the total score, but it was divergently valid with its subscale. Additionally, this subscale was divergently valid with all the other subscales. Therefore, the whole subscale (27–33) was deleted from the questionnaire.

In summary, items 5, 6, 13, 21, 27, 28, 29, 30, 31, 32, and 33 were removed from the proposed questionnaire. All of the items and statistical values are shown in Columns 5, 6, and 7 in Table 1.

Discussion

Inclusive education in Bangladesh is still in the preliminary stages of development, and most GPS teachers are not familiar with its concepts. Moreover, GPS teachers are not yet required to undertake training in inclusive education. Therefore, it was difficult to gather information from teachers about their training needs. This study used a literature review to extract GPS teachers' training needs for inclusive education. In addition, using Likert scale tools is much more effective for collecting data from GPS teachers than open-ended questionnaires or interviews. Therefore, the current questionnaire was developed to comprise 5-point Likert scale questions and objective questions.

Bangladesh has been conducting studies on inclusive education for more than a decade; most of the studies are related to teachers' attitudes and efficacies. In the current study, most of the items were adapted from previous studies. Moreover, some items were generated according to the findings of the literature review. From the 18 training needs identified, 48 items were organized into nine objectives with nine subscales within them.

The subscales related to teachers' professional knowledge, practice, and engagement. Among the nine subscales, two measured GPS teachers' attitudes toward inclusive education. One subscale related to using school resources to implement

inclusive education, but it had no strong correlation with the other eight subscales. This finding proves that GPS teachers do not have much authority when using or managing school resources. Therefore, the subscale was removed from the questionnaire. Eleven items did not fulfill the validity and reliability requirements. Therefore those 11 items were omitted, including one subscale. Eventually, 37 items were accepted for the questionnaire, and those were categorized into eight subscales.

Items 5 and 6 were deleted for low correlation with the total score, although these items were convergently valid as a subscale. Kumar (2016) suggests that if any items have a correlation of < 0.30 with the total scale score, this proves low construct validity. Moreover, item 21 was divergently valid with both the total scale and the subscale. Therefore, these items were removed from the questionnaire. Item 13 had low communality in responses indicating that this item was not clear to respondents; therefore, it was deleted (Field, 2009).

Limitations and Future Work

The questionnaire items introduced in this study were selected from the literature review. Therefore, teachers' training needs were identified from the researchers' perspective and not directly from GPS teachers or their teacher trainers. In the pilot study, respondents numbered only 46 GPS teachers. For this reason, it was not possible to analyze factors between the current study's questionnaire items. In further research, factor analysis could be conducted after collecting a large quantity of data to see the extent of the sense of belonging factor in the current study's questionnaire items.

References

- Ahsan, M. T., Deppeler, J. M., & Sharma, U. (2013). Predicting pre-service teachers' preparedness for inclusive education: Bangladeshi pre-service teachers' attitudes and perceived teaching-efficacy for inclusive education. *Cambridge Journal of Education*, 43(4), 517–535. <https://doi.org/10.1080/0305764X.2013.834036>

- Ahsan, M. T., Sharma, U., & Deppeler, J. M. (2012). Exploring pre-service teachers' perceived teaching-efficacy, attitudes, and concerns about inclusive education in Bangladesh. *International Journal of Whole Schooling*, 8(2), 1–20. <https://doi.org/EJ975715>
- Ahsan, M. T., Tasnuba, T., Akter, M., Islam, M. K., & Miah, A. S. M. J. (2016). *Teacher education programme in Bangladesh: Inclusiveness for children with disabilities*. Dhaka: ActionAid.
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Bolarinwa, O. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science research. *Nigerian Postgraduate Medical Journal*, 22(4), 195–201. <https://doi.org/10.4103/1117-1936.173959>
- Bryman, A., & Cramer, D. (2004). Quantitative data analysis with SPSS 12 and 13: A guide for social scientists. London: Routledge.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- DPE (Directorate of Primary Education) (2014). *Annual primary school census 2014*. Dhaka: DEP http://dpe.portal.gov.bd/sites/default/files/files/dpe.portal.gov.bd/publications/b490814d_522e_4f81_b0a0_8f8727feab27/Final APSC 2016.pdf
- DeVellis, R. (1991). Scale development: Theory and applications. Newbury Park: SAGE
- Dyson, M., Plunkett, M., & McCluskey, K. (2018). Australian professional standards for teachers. In *Success in Professional Experience* (pp. 229–234). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108550703.014>
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed). London: SAGE. <https://doi.org/10.5555/2502692>
- Forlin, C., Earle, C., Loreman, T., & Sharma, U. (2011). The Sentiments, Attitudes, and Concerns about Inclusive Education Revised (SACIE-R) scale for measuring pre-service teachers' perceptions about inclusion. *Exceptionality Education International*, 21(3), 50–65.
- IBM (2017). *IBM SPSS Statistics 25.0*. Armonk, NY: IBM.
- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, 65, 2276–2284.
- Kumar, A. (2016). Review of the steps for development of quantitative research tools. *Advanced Practices in Nursing*, 01(01), 1–4. <https://doi.org/10.4172/2573-0347.1000103>
- Kudsi, Z., Fenlon, M. R., Hill, K., & Baysan, A. (2020). Developing a questionnaire to measure psychological disturbance associated with tooth loss. *Journal of Dentistry*, 98(April), 103353–103360, 103353. <https://doi.org/10.1016/j.jdent.2020.103353>
- Malak, M. S. (2013). Inclusive education in Bangladesh: Are pre-service teachers ready to accept students with special educational needs in regular classes? *Disability, CBR & Inclusive Development*, 24(1), 56–81. <https://doi.org/10.5463/dcid.v24i1.191>
- Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica: Revista Semestral Do Departamento de Psicologia da UFRGS*, 30(1), 3. <https://doi.org/10.1186/s41155-016-0057-1>
- Oluwatayo, J. A. (2012). Validity and reliability issues in educational research. *Journal of Educational and Social Research*, 2(May), 391–400. <https://doi.org/10.5901/jesr.2012.v2n2.391>
- Parsian, N., & Dunning, T. (2009). Developing and Validating a Questionnaire to Measure Spirituality: A Psychometric Process. *Global Journal of Health Science*, 1(1), 1–11. <https://doi.org/10.5539/gjhs.v1n1p2>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Roberts, P., & Priest, H. (2006). Reliability and validity in research. *Nursing Standard (Royal College of Nursing (Great Britain): 1987)*, 20(44), 41–45. <https://doi.org/10.7748/ns2006.07.20.44.41.c6560>
- Sangoseni, O., Hellman, M., & Hill, C. (2013).

- Development and validation of a questionnaire to assess the effect of online learning on behaviors, attitudes, and clinical practices of physical therapists in the United States regarding evidenced-based clinical practice. *The Internet Journal of Allied Health Sciences and Practice*, 11(2), 1-12
- Sharma, U., & Desai, I. (2002). Measuring concerns about integrated education in India. *The Asia-Pacific Journal on Disabilities*, 5(1), 2-14.
- Sharma, U., Loreman, T., & Forlin, C. (2012). Measuring teacher efficacy to implement inclusive practices. *Journal of Research in Special Educational Needs*, 12(1), 12-21.
- Siddik, M. A. B., & Kawai, N. (2018). Preparing primary level teachers for inclusive education in Bangladesh. *Asian Journal for Inclusive Education*, 6(1), 49-71.
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford: Oxford University Press.
- Torkian, S., Shahesmaeili, A., Malekmohammadi, N., & Khosravi, V. (2020). Content validity and test-retest reliability of a questionnaire to measure virtual social network addiction among students. *International Journal of High Risk Behaviors and Addiction*, 9(1), 1-5. <https://doi.org/10.5812/IJHRBA.92353>
- UNICEF (United Nations Children's Fund) (2014). *Factsheet: Diploma in Primary Education*. March. https://www.unicef.org/bangladesh/DPEd_Fact_Sheet_v1.2_140203.pdf
- Younas, A., & Porr, C. (2018). A step-by-step approach to developing scales for survey research. *Nurse Researcher*, 26(3), 14-19. <https://doi.org/10.7748/nr.2018.e1585>

(2021. 2. 1受理)