Research in Autism Spectrum Disorders 57 (2019) 28–34

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

Brief report

Examining the psychometric properties of the autism stigma and knowledge questionnaire (ASK-Q) in multiple contexts

Ashley J. Harrison[a,*], Madison L. Paff[a], Marilyn S. Kaff[b]

[a] The University of Georgia, Athens, GA, United States
[b] Kansas State University, Manhattan, KS, United States

ARTICLE INFO

Keywords:
ASD knowledge
Autism stigma & knowledge questionnaire (ASK-Q)
Measurement
Cross-cultural

ABSTRACT

Background: The Autism Stigma & Knowledge Questionnaire (ASK-Q) is a new tool, developed to overcome many of the measurement issues associated with previously used ASD knowledge tools (Harrison, Slane, et al., 2017). The ASK-Q was developed to have cross-cultural utility and initial data reveals strong psychometric properties for the measure (Harrison, Bradshaw, et al., 2017). The current study aims to confirm several important measurement aspects of the ASK-Q.

Methods: To assess the cross-cultural internal consistency of the ASK-Q, data was collected from parents of children with ASD in Mongolia (n = 40). Test-retest reliability analyses among a group of college students (n = 110) examined the stability of responses on the ASK-Q. To examine the sensitivity of the ASK-Q to detect change, we examined differences in ASD knowledge from the beginning to the end of the semester for college students (n = 27) enrolled in a brief weekly ASD seminar.

Results: Analyses revealed adequate ASK-Q internal consistency in the Mongolian context with an alpha of 0.721. Test-retest data revealed good reliability for the ASK-Q overall (ICC = 0.86) over a two-week period, and the measure served as a useful tool for detecting change pre- and post-intervention.

Discussion: The adequate internal consistency result from the Mongolian context supports the ASK-Q development goal of creating a measure with cross-cultural utility. The additional psychometric data collected in this study reinforce the assertion that the ASK-Q would serve as a reliable tool and a sensitive tool for examining the efficacy of ASD knowledge interventions.

1. Introduction

Substantial variability exists with regard to autism spectrum disorder (ASD) service provision among minority groups in the United States (Bernier, Mao, & Yen, 2010) and around the world (Khan et al., 2012). Service disparities observed among different minority groups locally and in many low- and middle-income countries are driven, in part, by limited knowledge of ASD, which results in delays in assessment and treatment seeking and contributes to the proliferation of stigmas (Khan et al., 2012; Magaña, Lopez, Aguinaga, & Morton, 2013). Given the documented importance of ASD knowledge in early assessment and treatment provision (Bryson, Rogers, & Fombonne, 2003), an increasing amount of research has focused on psychoeducational interventions designed to increase ASD knowledge among parents and other target populations (for a review see Schultz, Schmidt, & Stichter, 2011). Interventions with a psychoeducation component have the potential to make meaningful differences in early identification and early

* Corresponding author at: University of Georgia, 110 Carlton Street, Athens, GA 30606, United States.
E-mail address: Ashley.harrison@uga.edu (A.J. Harrison).

https://doi.org/10.1016/j.rasd.2018.10.002
Received 1 March 2018; Received in revised form 1 October 2018; Accepted 4 October 2018
Available online 15 October 2018
1750-9467/ Published by Elsevier Ltd.

intervention; however, there is minimal research in this area (Bearss, Burrell, Stewart, & Scahill, 2015) and inadequate study designs, including insufficient outcome measures, limit the generalizability of research examining the efficacy of these intervention approaches (Schreibman, 2000).

A recent review revealed that as of 2015 the construct of ASD knowledge had been assessed in as many as 67 studies across 21 countries (Harrison, Slane, Hoang, & Campbell, 2017). This review highlighted multiple concerns with many of the existing approaches used to quantify the construct, including most commonly used ASD knowledge instruments such as the Autism Knowledge Survey (AKS; Stone, 1987) and subsequent revisions (i.e., Bauer, Carroll, Saha, & Downs, 2015; Swiezy, 2007) and the Autism Knowledge Questionnaire (AKQ; Kuhn & Carter, 2006). At the time of the review, the majority of ASD knowledge tools had measurement limitations such as not meeting standards for consideration as well-established due to concerns with the psychometric properties and limited replicated examinations of measurement properties (Harrison, Slane et al., 2017). Further, many ASD knowledge measures had not been designed with cross-cultural research in mind and tended to emphasize a broad assessment of ASD knowledge rather than more thoroughly assessing knowledge subdomains (Harrison, Slane et al., 2017). In a review of the literature since November 2015, 14 articles have been published quantitatively assessing ASD knowledge. Of the 14 articles published, 11 studies used previously created measures with limited updated psychometric support and 2 studies created measures for the purposes of their study with similar psychometric limitations as mentioned in the previous review. There has been one study that developed a measure (The ASD Knowledge and Self-Efficacy Questionnaire), using The Autism Survey (Stone, 1987) as its basis, with solidly tested psychometric properties, but the psychometric properties for this measure were tested with the measure after it had been translated into Hebrew (Atun-Einy & Ben-Sasson, 2018).

A new instrument, the Autism Sigma and Knowledge Questionnaire (ASK-Q) was developed as a potential solution to these documented measurement concerns to be used with a wide audience (Harrison, Bradshaw, Naqvi, Paff, & Campbell, 2017). The ASK-Q has 49 items and a four-factor structure to assess three ASD knowledge domains of Symptoms/Diagnosis (18 items), Etiology (16 items), and Treatment (14 items), as well a fourth dimension of ASD stigma endorsement (7 items) comprised of items also included in one other knowledge subdomain. Previous research, using a newer psychometric analysis technique called Diagnostic Classification Model (DCM) analysis (Rupp, Templin, & Henson, 2010), provided initial support for the proposed four-factor structure, item validity, and a form of test-retest reliability. Initial testing of the ASK-Q in a U.S. sample revealed high internal consistency and allowed for the determination of cutoff scores to differentiate those with adequate knowledge from those without (Harrison, Bradshaw et al., 2017). The ASK-Q was developed with cross-cultural utility as a central aim. Specifically, each item included in the measure was evaluated by a team of international researchers to make sure it had cross-cultural validity to minimize the need for subsequent question adaptation in different cultures and to ensure aspects of ASD beliefs that vary across cultures were sufficiently represented (for more details see Harrison, Bradshaw et al., 2017). With strong psychometric support and cross-cultural utility, the ASK-Q has the potential to serve as an ideal outcome measure for research and practice designed to improve ASD knowledge and may allow researchers to identify specific regions demonstrating deficits in ASD knowledge around the world.

In spite of the initial documented strengths of the ASK-Q (Harrison, Bradshaw et al., 2017), several important aspects of the measure remain untested. Specifically, although this measure appears to fill the need for a strong ASD knowledge outcome measure, the test-retest reliability of the measure has not been examined over time. Measurement experts cite that ensuring stability in responses over time through test-retest is one of the most important metrics of reliability (Paiva et al., 2014) and an essential component of examining measurement error (Hopkins, 2000). In a similar vein, claims about the utility of the ASK-Q would be strengthened with published research demonstrating the sensitivity of this measure to detect change in ASD knowledge. Finally, while the measurement development process emphasized designing a measure with a certain level of cross-cultural neutrality, the reliability of the final version of the ASK-Q has yet to be investigated in a cross-cultural context. Although a thorough investigation of cross-cultural validity to examine the success of developing a measure with cultural equivalency will involve testing in multiple cultures (including a context that requires language translation), this study is designed to be an initial examination of cross cultural validity in a different country in the same language to determine if adaptation steps are warranted (Beaton, Bombardier, Guillemin, & Ferraz, 2000). The current paper aims to provide additional psychometric support for the ASK-Q in a series of studies by 1) testing the test-retest of the ASK-Q, 2) examining the sensitivity of the ASK-Q to detect change, and 3) examining the internal consistency in a cross-cultural context (Mongolia).

2. Method

2.1. Procedure overview

Additional psychometric support for the Autism Stigma & Knowledge Questionnaire (ASK-Q) was obtained through three studies involving independent data collection. Given the potential for widely using the ASK-Q among groups of distinct individuals and across diverse settings, these psychometric studies of the ASK-Q sampled the lay public as well as parents of children with ASD. Individuals participating in each of these studies provided university approved informed consent.

2.2. Assessment measure

All three studies examined the overall ASK-Q score as well as subscale scores for each of the three ASD knowledge subscales (Diagnosis/Symptoms, Etiology, Treatment) and the ASD stigma endorsement subscale. Of note, all items in the ASD Stigma subscale are considered complex and also load onto an ASD knowledge subscale and represent beliefs about ASD that are also indicative of

stigmatizing ideation. Each correct response results in one point, and for each subscale, the following cutoff score ranges reflect adequate knowledge: Diagnosis/Symptoms 11–18, Etiology 11–16, and Treatment 10-14. For the complex items that assess both ASD knowledge and stigma, a correct knowledge response would mean that they were not endorsing stigma, therefore, the stigma endorsement subdomain was reverse scored such that higher scores (3–7) demonstrate a failure to endorse stigma and lower scores (0–2), indicate stigma endorsement.

Items selected for inclusion on the ASK-Q demonstrated good face, construct, and cross-cultural validity. In addition, as previously mentioned, an initial measurement examination of the ASK-Q confirmed the four-factor structure of this measure and revealed high internal consistency (Cronbach's alpha = 0.88; Harrison, Bradshaw et al., 2017). A reliability metric similar to test-retest derived from Diagnostic Classification Model (DCM) analysis (Rupp et al., 2010) showed strong reliability for each of the four subscales (range: 0.933 – 0.984).

2.3. Cross-cultural reliability

To specifically examine the reliability of the ASK-Q in a different cultural context, the first study piloted the ASK-Q with 40 English-speaking parents of children with ASD in Ulaanbaatar, Mongolia. Ulaanbaatar is the capital of Mongolia, a country with a population estimated to be over 3,000,000 with about 21.6% of the population living below the poverty line as of 2014 (The World Bank, 2014; Worldometers, 2018). The English speaking population in Mongolian has been on the rise steadily since the government mandated that all public school students study English from 4th to 11th grade in 2005 (Cohen, 2004). Very limited research on ASD has been conducted in Mongolia, but the Autism Research Society of Mongolia estimates that approximately 18,000 individuals have an autism spectrum disorder in Mongolia ("18,000 Mongolians may be affected by autism," 2018).

The mean age of children was 4.73 years old (range: 1 to 14 years old) and 70 percent were boys. Parents/caregivers came to a center for children with ASD in Ulaanbaatar to participate in ASD diagnostic evaluations and parent training workshops completed by a professor from the U.S. with ASD expertise. ASD diagnoses were confirmed through a diagnostic approach that aligned with Harrison, Zimak, Sheinkopf, Manji, and Morrow, (2014). Although Mongolian was the native language for all parents/caregivers, they were all proficient in English. Primary caregivers were invited to also complete a paper and pencil version of the ASK-Q in the original English format. The majority of parents (75%) had obtained college degrees and held professional occupations, while a minority of participating parents (25%) reported having working class jobs and no college education. Families in the sample came from diverse economic backgrounds in alignment with the general Mongolian population. See Table 2 for additional demographic data regarding family structure. Data collected in this context allowed for the examination of the cross-cultural internal consistency of the ASK-Q in one specific Mongolian context.

2.4. Test-retest reliability

The second research study completed examined the stability of responses on the ASK-Q among college students participating in an educational psychology research pool in the United States. To allow for an examination of test-retest reliability, 110 undergraduate students completed the measure twice, two weeks apart. This time interval has a reasonable support in the literature to suggest sufficient time to both minimize recollection bias and minimize uncontrolled change in the underlying construct (Marx, Menezes, Horovitz, Jones, & Warren, 2003). Based on power calculations for simple test-retest reliability research, this was estimated to be a sufficient size sample to provide adequate precision (Hopkins, 2000). Students earned research credits for participating in this study that fulfilled a requirement in an undergraduate course that necessitates participating in ongoing university research to learn about this aspect of the scientific method. Participants in the test-retest study completed the measure online.

Participants in this study were derived from a typically matriculating sample at a large Southeastern University in the United States. All but 1 participant was in the 18–25-year-old age range. Of the sample, 86% identified as female, 9% as male, and 2% reported a different classification or preferred not to answer. The sample was 87% White, 8% Asian, and 5% Black or African American. Students indicated 30 different majors with the majority from the College of Education and reported variable class status (19% Freshman, 53% Sophomores, 20% Juniors, 6% Seniors; see also Table 2).

2.5. Sensitivity to detect change

In a third study, in order to examine the sensitivity of the ASK-Q to detect change, we measured differences in ASD knowledge from the beginning to the end of the semester for college students enrolled in a brief Freshman ASD seminar at a large Southeastern University in the United States for students from a range of backgrounds. The seminar, led by a professor specializing in ASD, met 15 weeks for 50 min per week and used multi-media sources to talk about a range of ASD topics selected by the students from any form of media. Students completed the ASK-Q as a pre-/post-test for the class in a paper pencil format in the first and last weeks of the course and were not allowed access to external sources of information during completion. After completion of the semester and blind to the professor to ensure minimal feelings of coercion, the students could volunteer to include their data as part of research. Across two Fall semesters, n = 27 students participated in the study (all students in the classes). Of the participating students, 78% were female. See Table 2 for additional demographic data.

Table 1
Descriptive statistics for all three studies compared to the original published data (Harrison, Bradshaw et al., 2017).

Study		Symptoms/ Diagnosis	Etiology	Treatment	Stigma	Total Score
Original Study (Harrison, Bradshaw et al., 2017)	Cutoff score (above = adequate knowledge/ below = no stigma)	10	10	9	3	–
	% with adequate knowledge/no stigma	72.90%	55.80%	56.87%	69.88%	–
Cross-cultural Reliability Study (n = 40)	Mean Scores (SD)	13.75 (1.24)	13.98 (1.59)	11.65 (0.89)	5.2 (0.99)	41.43 (3.58)
	% with adequate knowledge/no stigma	100%	100%	100%	100%	–
	Reliability: Internal Consistency	–	–	–	–	0.721
Test-retest: Time 1 (n = 110)	Mean Scores (SD)	10.32 (2.96)	10.51 (2.51)	9.39 (2.47)	4.68 (1.20)	32.06 (7.39)
	% of sample with adequate knowledge/no stigma endorsed	57.50%	49.10%	52.70%	2.70%	–
Test-retest: Time 2 (n = 110)	Mean Scores (SD)	10.70 (2.93)	10.48 (2.60)	9.41 (2.32)	4.07 (1.49)	32.52 (6.84)
	% with adequate knowledge/no stigma	57.50%	50.01%	52.70%	16.30%	–
	Reliability: ICC	0.78	0.83	0.82	0.55	0.86
Sensitivity to Detect Change: Time 1 (n = 27)	Mean Scores (SD)	14.37 (2.62)	12.26 (2.33)	11.19 (1.94)	0.93 (1.07)	37.81 (5.73)
	% with adequate knowledge/no stigma	96.30%	77.80%	85.20%	96.30%	–
Sensitivity to Detect Change: Time 2 (n = 27)	Mean Scores (SD)	17.07 (2.32)*	14.04 (1.13)*	12.41 (1.22)*	0.52 (0.64)*	43.52 (2.87)*
	% with adequate knowledge/no stigma	100%	100%	100%	100%	–

* Indicates a significant difference at a level of $p < .05$.

3. Results

3.1. Cross-cultural reliability

Analyses revealed adequate (George & Mallery, 2003; Nunnally, 1978) internal consistency for the ASK-Q in the Mongolian context with an alpha of 0.721. In this context, similar to the initial sample in the United States, mean ASK-Q scores align with the cutoffs demonstrating adequate diagnostic knowledge ($M = 13.75$), adequate etiological knowledge ($M = 13.98$), adequate treatment knowledge ($M = 11.65$), and limited stigma endorsement ($M = 5.20$). However, this was a particularly knowledgeable sample with 100% of participants demonstrating adequate knowledge across all of the subscales and 100% showing no endorsement of stigmas (Table 1).

3.2. Test-retest reliability

Intraclass correlation coefficients (ICC) were calculated for the overall measure as well as the four subscales to assess test-retest reliability. A two-way mixed model with a single-measurement absolute agreement was used to determine the ICC for the overall ASK-Q and subscales. The ICC for the ASK-Q overall was 0.86 with a 95% confidence interval (CI) from 0.80 to 0.91, which falls in the good to excellent range according to standards published by Koo and Li (2016). The Etiology subscale had the highest ICC in the good range at 0.83 (95% CI: 0.75–0.88), followed by the Treatment subscale at 0.82 (95% CI: 0.74– 0.88) in the moderate to good range, the Symptoms/Diagnosis subscale at 0.78 (95% CI: 0.68–0.85) in the moderate to good range, and finally the Stigma subscale at 0.61 (95% CI: 0.43–0.73) in the poor to moderate range. Due to the low ICC for the Stigma subscale, we conducted a follow up examination of each of the seven items in this subscale in spite of the small sample size as an exploratory investigation. Three of the seven items had an ICC below .70 (Item 15: ICC = .46, Item 24: ICC = .55, and Item 27: ICC = .56). See Table 1 for more descriptive information.

3.3. Sensitivity to detect change

To examine sensitivity to detect change, a series of Paired Samples *t*-tests were used to compare student ASK-Q scores from the beginning of the semester to scores at the end of the semester, following the weekly ASD seminar. Analyses examined the overall knowledge score obtained on the ASK-Q as well as each of the three subscales examining specific aspects of ASD knowledge and the stigma subscale. Significant group differences were detected between time 1 and time 2 on log transformed variables to account for skewness for total ASK-Q score $t(26) = -4.86$, $p < .001$, $d = .06$, and for all of the ASK-Q subscales: Diagnosis, $t(26) = -4.21$, $p < .001$, $d = .13$, Etiology, $t(26) = -4.52$, $p < .001$, $d = .01$, Treatment, $t(26) = -3.09$, $p = .003$, $d = .01$, and Stigma, $t(26) = 2.28$, $p = .03$, $d = .45$. See Table 1 for more descriptive information.

To contextualize the results, additional analyses reveal that at the first time point, 96% of participants had adequate knowledge of Diagnosis/Symptoms, 78% of etiology, 85% of treatment and that only 4% of students endorsed stigmas. At the end of the semester (time 2), 100% of participants had adequate knowledge of Diagnosis/Symptoms, 100% of etiology, 100% of treatment and no

Table 2
Demographic data for the three included studies.

Study	Demographic Factor	Percentage
Cross-cultural reliability (n = 40)	Parent Ethnicity	
	Mongolian	100
	Parent Education	
	College Degree with Professional Occupation	75
	No College Education with Working Class Job	25
	Child Sex	
	Male	70
	Female	30
	Child Age	
	1-14 years	100
	Reporter	
	Mother	77
	Father	18
	Grandparent	5
	Mother	
	Child Sibling Status	
	No Sibling	15
	One Sibling	42.5
	Two Siblings	27.5
	Three Siblings	15
	Living Situation	
	Two-Parent Homes	62.5
	One-Parent Homes	32.5
	Grandparents	5
Test-retest (n = 110)	Sex	
	Female	86
	Male	9
	Other or Prefer Not to Answer	2
	Race	
	White	87
	Asian	8
	Black or African American	5
	Age	
	18–25	99
	Other	1
	Class Status	
	Freshman	19
	Sophomore	53
	Junior	20
	Senior	6
Sensitivity to detect change (n = 27)	Sex	
	Female	78
	Class Status	
	Freshman	100

students endorsed stigmas as measured by the ASK-Q. These results emphasize that in spite of relatively high initial levels of ASD knowledge, the ASK-Q had the sensitivity to detect small amounts of change. To further examine the magnitude of the change a reliable change index (RCI) was calculated for the group of participants as well as for each of the participants (Bauer, Lambert, & Nielsen, 2004; Jacobson & Truax, 1991; Jacobson, Follette, & Revenstorf, 1984). Using the average of all the participants, the group RCI was greater than 1.96, indicating a reliable average group improvement. Of the 27 participants in this study, 25 participants showed an increase in performance from their pre-training performance to their post-training performance and 12 of those had an RCI greater than 1.96, meaning 44.4% of the total number of participants in this study reliably improved.

4. Discussion

This series of studies provides important additional psychometric support for the Autism Stigma and Knowledge Questionnaire (ASK-Q). The adequate internal consistency result observed within the Mongolian respondents offers initial data from a distinct international context to support the ASK-Q development goal of creating a measure with cross-cultural utility. Additionally, the ASK-Q was sensitive enough to detect a change in ASD knowledge within a small sample with a modest intervention. Finally, in a reasonably large sample, the total score and all knowledge subscales of the ASK-Q demonstrated moderate to excellent test-retest reliability over a two-week period with the exception of the stigma subscale. These studies provide additional psychometric data to bolster the data provided in the initial development article (Harrison, Bradshaw et al., 2017) and further speaks to the utility of using this as an outcome measure to examine intervention efficacy.

The inclusion of three distinct studies conducted with different populations functions as a strength of this further investigation of the psychometric properties of the ASK-Q; however, several noteworthy limitations exist. First, while the updated psychometric results of the ASK-Q were overwhelmingly positive, the low test-retest reliability score on the Stigma subscale warrants note. Given that three of the seven items in this subscale demonstrated an ICC below 0.70 in exploratory analyses, this subscale should be interpreted with some caution and future research may consider adding more items assessing stigma endorsement. Additionally, this study used a university convenience sample, which may limit the generalizability of findings (Sears, 1986). Second, although the data collected in Mongolia provides initial evidence in support of the cross-cultural utility of the ASK-Q, this was by no means a comprehensive examination of cross-cultural validity as data collection was limited to a relatively small, homogeneous sample of parents in only one distinct context. As noted in the results, this sample demonstrated particularly high knowledge and may not represent the Mongolian population at large. Additional data collection across a range of cultural contexts in alignment with published standards for measurement translation and adaptation is necessary before making claims about the cross-cultural validity of the ASK-Q (Beaton et al., 2000). Currently, projects using the ASK-Q are underway in additional countries in Asia, the Middle East, Europe, and Africa. That said, the current data from Mongolia in combination with a measurement design approach intended to minimize cultural specificity provides promise for ASK-Q cross-cultural utility.

Sample size and selection methods must be noted as additional limitations. Although the test-retest sample was reasonably sized (Hopkins, 2000), the samples examining the sensitivity to detect change and cross-cultural relatability were notably smaller. Additional research testing both aspects of reliability and clinical utility in larger samples and more diverse contexts will help to further bolster the claims made about the psychometric strength of the measure. Given the small sample sizes it was difficult to examine subscale and item level reliability. This should also be a focus of future research with a larger sample size.

In support of the ASK-Q as a new, robust measure to quantify ASD knowledge, findings from these three studies provide more data to highlight additional areas of psychometric strength for the ASK-Q. This data confirms test-retest reliability using stronger methods than were previously employed, provides positive evidence in support of the cross-cultural validity assumption outlined in the original manuscript (Harrison, Bradshaw et al., 2017), and demonstrates the sensitivity of the ASK-Q to detect small amounts of change.

Conflict of interest

None of the authors have conflicts of interest of any kind to report.

Acknowledgment

The authors would like to acknowledge Sydney Erickson for her help with data collection for this manuscript.

References

Atun-Einy, O., & Ben-Sasson, A. (2018). Pediatric allied healthcare professionals' knowledge and self-efficacy regarding ASD. *Research in Autism Spectrum Disorders, 47*, 1–13. https://doi.org/10.1016/j.rasd.2017.12.001.

Bauer, N. S., Carroll, A. E., Saha, C., & Downs, S. M. (2015). Computer decision support changes physician practice but not knowledge regarding autism spectrum disorders. *Appl Clin Inform, 6*(3), 454–465. https://doi.org/10.4338/aci-2014-09-ra-0084.

Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment, 82*(1), 60–70. https://doi.org/10.1207/s15327752jpa8201_11.

Bearss, K., Burrell, T. L., Stewart, L., & Scahill, L. (2015). Parent training in autism spectrum disorder: What's in a name? *Clinical Child and Family Psychology Review, 18*(2), 170–182. https://doi.org/10.1007/s10567-015-0179-5.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014.

Bernier, R., Mao, A., & Yen, J. (2010). Psychopathology, families, and culture: Autism. *Child and Adolescent Psychiatric Clinics of North America, 19*(4), 855–867. https://doi.org/10.1016/j.chc.2010.07.005.

Bryson, S. E., Rogers, S. J., & Fombonne, E. (2003). Autism spectrum disorders: Early detection, intervention, education, and psychopharmacological management. *The Canadian Journal of Psychiatry, 48*(8), 506–516. https://doi.org/10.1177/070674370304800802.

Cohen, R. (2004). The current status of English education in Mongolia. *Asian EFL Journal, 6*(4), 1–21.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston, MA.

Harrison, A. J., Zimak, E. H., Sheinkopf, S. J., Manji, K. P., & Morrow, E. M. (2014). Observation-centered approach to ASD assessment in Tanzania. *Intellectual and Developmental Disabilities, 52*(5), 330–347. https://doi.org/10.1352/1934-9556-52.5.330.

Harrison, A. J., Bradshaw, L. P., Naqvi, N. C., Paff, M. L., & Campbell, J. M. (2017). Development and psychometric evaluation of the autism stigma and knowledge questionnaire (ASK-Q). *Journal of Autism and Developmental Disorders, 47*(10), 3281–3295. https://doi.org/10.1007/s1080.

Harrison, A. J., Slane, M. M., Hoang, L., & Campbell, J. M. (2017). An International review of autism knowledge assessment measures. *Autism, 21*(3), 262–275. https://doi.org/10.1177/1362361316638786.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–1219. https://doi.org/10.1037/0022-006X.59.1.12.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7.

Khan, N. Z., Gallo, L. A., Arghir, A., Budisteanu, B., Budisteanu, M., Dobrescu, I., ... Elsabbagh, M. (2012). Autism and the grand challenges in global mental health. *Autism Res, 5*(3), 156–159. https://doi.org/10.1002/aur.1239.

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.

Kuhn, J. C., & Carter, A. S. (2006). Maternal self-efficacy and associated parenting cognitions among mothers of children with autism. *American Journal of Orthopsychiatry, 76*(4), 564–575. https://doi.org/10.1037/0002-9432.76.4.564.

Magaña, S., Lopez, K., Aguinaga, A., & Morton, H. (2013). Access to diagnosis and treatment services among latino children with autism spectrum disorders. *Intellectual and Developmental Disabilities, 51*(3), 141–153. https://doi.org/10.1352/1934-9556-51.3.141.

Marx, R. G., Menezes, A., Horovitz, L., Jones, E. C., & Warren, R. F. (2003). A comparison of two time intervals for test-retest reliability of health status instruments. *Journal of Clinical Epidemiology, 56*(8), 730–735. https://doi.org/10.1016/S0895-4356(03)00084-2.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY.

Paiva, C. E., Barroso, E. M., Carneseca, E. C., de Pádua Souza, C., dos Santos, F. T., Mendoza López, R. V., ... Ribeiro Paiva, S. B. (2014). A critical analysis of test-retest reliability in instrument validation studies of cancer patients under palliative care: A systematic review. *BMC Medical Research Methodology, 14*(1), 8. https://doi.org/10.1186/1471-2288-14-8.

Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York, NY: Guilford.

Schreibman, L. (2000). Intensive behavioral/psychoeducational treatments for autism: Research needs and future directions. *Journal of Autism and Developmental Disorders, 30*(5), 373–378.

Schultz, T. R., Schmidt, C. T., & Stichter, J. P. (2011). A review of parent education programs for parents of children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 26*(2), 96–104.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51*(3), 515–530.

Stone, W. L. (1987). Cross disciplinary perspectives on autism. *Journal of Pediatric Psychology, 12*, 615–630. https://doi.org/10.1093/jpepsy/12.4.615.

Swiezy, N. A. (2007). *Autism knowledge questionnaire-revised. HANDS in autism.*. Christian Sarkine Autism Treatment Center.

The World Bank (2014). *Poverty headcount ratio at national poverty lines (% of population): Mongolia.* Retrieved fromhttps://data.worldbank.org/indicator/SI.POV.NAHC?locations=MN.

Worldometers (2018). *Mongolian population (live).* Retrieved fromhttp://www.worldometers.info/world-population/mongolia-population/.