


Development and Content Validation of a Measure to Assess the Parent-Child Social-emotional Reciprocity of Children with ASD

Asmita Nayak¹  and Rooplekha Khuntia¹ 

ABSTRACT

Background: As no scales are currently available to assess the social-emotional reciprocity (SER) between children with autism spectrum disorder (ASD) and their parents, we aimed to develop a questionnaire for this.

Methods: Both quantitative and qualitative techniques were used for content validation. All stages of this study used purposive sampling to choose various groups of experts, mental health professionals, and parents of children with ASD to participate in the development, judgment-quantifying, and pretest stages. Data from 30 parents of children with ASD were gathered for the field test.

Results: Thirty items were retained after content validation. The proportion of consensus regarding the comprehensiveness of the entire questionnaire was 1.00. The scale-level content validity index (CVI) by averaging calculation method and universal agreement calculation method were 0.95 and 0.50, respectively. Following the

pretest, a total of 23 questions, on a five-point scale, were retained.

Conclusion: Given the significance of SER between children with ASD and their parents, mental health practitioners should use this questionnaire to understand it to maximize the efficacy of any intervention.

Keywords: Autism spectrum disorder, Social-emotional reciprocity, Family, Content validity

Key Messages: Understanding the SER of children with ASD with their parents will help design a more effective intervention program using this questionnaire that has sufficient content validity.

Autism spectrum disorder (ASD) causes persistent deficits in social communication and social interaction and restricted, repetitive patterns of behavior, interests, or activities.¹ Early intervention improves the symptoms of children with ASD and facilitates their integration into society, and early identification is crucial to get them the needed help.²⁻⁴ An enriching social milieu also

reduces the severity of the symptoms. Along with early diagnosis, awareness of family dynamics may aid in designing a more effective therapeutic program. As children spend most of their formative years at home, understanding their social communication and social interaction with parents is critical. Alert and careful parents will be able to take preventive measures or handle situations effectively. As a major dimension of social communication and social interaction in ASD,¹ understanding social-emotional reciprocity (SER) will be beneficial for intervention purposes. Additionally, knowing how children with ASD interact with their parents will provide the interventionist with a holistic perspective that can help them design a more effective intervention module. So, a valid, more accessible, and easy-to-administer tool that provides information on children's SER with their parents is required. To our knowledge, no such independent measure is available.

¹National Institute of Science Education and Research, An OCC of Homi Bhabha National Institute, Bhubaneswar, Odisha, India.

HOW TO CITE THIS ARTICLE: Nayak A and Khuntia R. Development and Content Validation of a Measure to Assess the Parent-Child Social-emotional Reciprocity of Children with ASD. *Indian J Psychol Med.* 2024;46(1):66–71.

Address for correspondence: Rooplekha Khuntia, National Institute of Science Education and Research, An OCC of Homi Bhabha National Institute, Bhubaneswar 752050, Odisha, India.
E-mail: rooplekha@niser.ac.in

Submitted: 23 Jan. 2023
Accepted: 27 Sep. 2023
Published Online: 08 Dec. 2023



Copyright © The Author(s) 2023

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution- NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-Commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the Sage and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

ACCESS THIS ARTICLE ONLINE
Website: journals.sagepub.com/home/szj
DOI: 10.1177/02537176231207763

Social interaction is an assessment component built into the available instruments like the INCLIN Diagnostic Tool for Autism Spectrum Disorder (INDT-ASD) and the Indian Scale for Assessment of Autism (ISAA). But they are available only to professionals.^{5,6} Moreover, the cost involved in using these scales, cross-cultural variances in the assessment construct, language, and test users' qualifications and training also pose barriers to their usability. To address these gaps, we developed a questionnaire that would be accessible to and easily administered by any mental health professional, not just specialists. This paper details the development and content validation of a questionnaire on the SER of children with ASD with their parents.

Methods

The study was carried out with the approval of the Institutional Ethics Committee (IEC No.: NISER/IEC/2018-08/V2). The research was conducted between December 2018 and February 2020. This

study used both quantitative and qualitative techniques. Data was collected from two urban hospitals. Informed consent was obtained from all the participants. The participants (**Table 1**) included one psychologist for peer review, a team of three professionals for assessing domain specification, nine experts for checking the content validity of the questionnaire, an expert for checking the structural aspects of the questionnaire, five test users for the pretest, five different test takers for each stage of the pretest, and 30 parents of children with ASD who had received INCLIN-based ASD diagnosis by mental health professionals and aged 3–10 years in the field test. Children with any comorbid conditions were excluded. The number of test takers and test users was kept minimal due to the limited availability of the target population and the fact that a part of the pretest was qualitative. For the field investigation, we aimed for a minimum sample size of 30 participants.

The content validation of the questionnaire included the development stage and the judgment-quantifying stage.

Development Stage

This had four steps.

Concept Definition and Formulation

Besides restricted, repetitive patterns of behavior, interests, or activities, ASD is characterized by “persistent deficits in social communication and social interaction across multiple contexts, as manifested by deficits in SER, nonverbal communicative behaviors used for social interaction, and developing, maintaining, and understanding relationships, currently or by history.”¹

Domain Identification

The SER dimension of Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5)'s broad category of social communication and social interaction served as the theoretical basis for this study's objective.^{1,7} An unbiased team of three professionals extensively reviewed the draft of the domain specification.⁸

Item Generation

A total of 72 items of the questionnaire (**Table S1**) were framed using four sources, including the first author's 2 years of clinical expertise, a literature review on the target behavior, and discussion with subject-matter specialists and various mental health professionals working on ASD.^{7,9} The items in the initial pool were generated based on the SER dimension of the broad domain, social communication, and social interaction of DSM-5.¹ A conscious effort was made to avoid past tense, double negatives, and ambiguous or complex sentence structure. The cultural context of the test takers was also considered while generating the items.

Questionnaire Formation

It encompassed all the processes required to check the representativeness and relevance of the elements.¹⁰ Items were examined for consistency, accuracy, and fitness for the SER dimension.

A five-point Likert scale format was adopted. The response categories and respective scores were (1) rarely, (2) sometimes, (3) frequently, (4) mostly, and (5) always. Negatively keyed items were reverse-scored. The scores on all the items were added to get the total score. A high total score would indicate better SER in the child. This was followed

TABLE 1.

Sociodemographic Details of Study Participants.

Participants	Sex		Age in years Mean and SD	Field of expertise	Place of work	Years of work experience with ASD (mean and SD)
	M	F				
Peer reviewer N = 01	0	01	NA	Psychologist	Research institute	8
Domain- specification reviewers N = 3	1	2	29.38 ± .55	Clinical Psychologists	Hospital	3.3 ± .58
Experts N = 9 + 1	5	5	46.60 ± 9.01	Psychiatrist/Clinical Psychologist/ Psychologist/ Researcher/Special educator/Parent	Hospital/ Educational and research institute/ Special school	17.22 ± 7.22
Test users N = 5	3	2	29.65 ± 1.51	Clinical psychologist, Psychologist, Social worker, Occupational therapist	Hospital	4.4 ± 1.34
Test takers (Pretest 1) N = 5	2	3	41.44 ± 2.49	Parents		
Test takers (Pretest 2) N = 5	2	3	40.85 ± 1.65	Parents		
Field-test participants N = 30	13	17	39.08 ± 3.14	Parents		

by drafting the instructions for the participants. Seventy-two items were forwarded for peer review to a psychologist who understood the concepts and identified repetitive items. The items retained in this process were then forwarded to experts for content validation.

Judgment-Quantifying Stage

It involved the assessment of the content validity of the individual items (I-CVI) and the complete questionnaire (S-CVI). This process focused on at least three features of the test items: (1) relevance of item content to the domain; (2) balance of coverage of the content domain; and (3) technical quality of the items, response formats, and scoring procedures.¹¹ A total of nine experts were selected to reduce the likelihood of a chance agreement. They were included based on their expertise in the content area and/or instrument development, having at least five years of working experience with children with ASD. They were contacted after defining the construct and at least one week before the commencement of the content validation procedure.^{12,13} To ensure the representation of the target population, a special educator who was also a parent of a child with ASD was one of the experts. In addition to these nine specialists, an expert in structural aspects of questionnaire construction exclusively examined the questionnaire development procedure, item structure, format, level of measurement (nominal, ordinal, interval, ratio scale), and fitness to the research objective.¹³ An information kit containing a cover letter outlining the study's purpose, conceptual and operational definitions of the construct, a brief description of the questionnaire and its scoring, and a description of the content validity form was sent to all 10 experts.¹² The instruction kit also consisted of a specimen copy of the developed questionnaire, the content validity assessment form (response form), and a brief demographic questionnaire. Only nine experts were instructed to emphasize the items' clarity, specificity, and grammatical structure; add or delete items; and make any additional suggestions. Congruency, readability, accuracy, language, and representativeness were all prioritized.⁸

First Judgment Stage

While measuring the content validity, in the first judgment stage, we used the approach of Lawshe.^{14,15} To prevent pushing them to indicate a specific tendency, the nine experts assessed the content validity ratio (CVR) on a three-point scale: (1) not essential, (2) useful but not essential, and (3) essential. The CVR formula presented by Lawshe¹⁴ is $\frac{n_e - N/2}{N/2}$. Here, n_e is the number of panellists who indicated "essential," and N represents the total number of panellists. Items with a CVR <0.78 were eliminated, and some were modified based on expert advice.

Second Judgment Stage

In this stage, the content validity index (CVI) was calculated using Lynn's technique.^{15,16} The same experts were asked to revise the content validity of those items retained in the first judgment stage.¹⁶ For reassessment, time intervals of 10–14 days were chosen.¹⁷ Following Lynn's¹⁶ suggestion, a four-point scale was used to assess the relevance, that is, (1) not relevant; (2) unable to assess relevance without item revision, or the item needs such revision that it would no longer be relevant; (3) relevant but needs minor alteration; and (4) very relevant and succinct. Clarity was also evaluated on a four-point scale. I-CVI was obtained by dividing the number of experts rating the item as 3 or 4 by the total number of experts.^{15,16} Items having an I-CVI of >0.78 were preserved. Calculating the average CVI across the items (summing the I-CVIs and dividing them by the number of items) did not yield the desired level of content validity. Hence, average-CVI was determined using a different approach, that is, calculating the average proportion of agreement across experts by dividing the total number of items to which an expert agreed with the total number of items. This helped determine if the same experts would be considered for the third round of judgment.¹⁵

Third Judgment Stage

Here, the same experts were requested to revise the content validity on a four-point scale.¹⁶ Time intervals of 10–14 days were employed to reassess the items' relevance and clarity.¹⁷ Items with a CVI

<0.78 were deleted. Only the experts' ratings on the relevance of items were considered to estimate the CVI of the questionnaire (S-CVI), ρ_c (probability of chance agreement), modified kappa (K), the total number of agreements for the comprehensiveness of the questionnaire, and the proportion of consensus for the comprehensiveness of the questionnaire.¹⁵⁻¹⁸ Both the Ave-CVI, that is, "CVI of the entire questionnaire, averaging calculation method," and the UA-CVI, that is, "CVI of the entire questionnaire, universal agreement calculation method," were used to calculate S-CVI.¹⁵ The r_c and K values were calculated as per the table given by Polit et al.¹⁹ ρ_c was calculated as $\rho_c^* = [n!/(r!(n-r)!)]0.5^n$, the formula for calculating Kappa was $K = (ICVI - \rho_c)/(1 - \rho_c)$.

Pretest

The revised measure was pretested for other psychometric properties and to determine the ease of administration. The sample comprised various test users and test takers, as described before,¹ who visited two urban hospitals for consultation.¹² The two hospitals are located in the state capital and its neighboring city. Due to the locational advantage, patients from all over the state visit these hospitals for consultation and treatment. Test users and test takers were requested to record their comments regarding the questionnaire and its administration procedure.

Stage I

This had five test users (mental health professionals in different hospitals), five test takers (parents of children with ASD), and 10 experts (same as on the content validity method). The test users conducted cognitive interviews using emergent probes, which were flexible, unscripted, and reactive.²⁰ They used the behavior coding process to identify paradigmatic, problematic, and inadequate items.²¹ The cultural sensitivity of the items; the test user's behavior; the test environment; fairness, language, and length of the questionnaire; guessing, integrity, and physical and mental status of the test takers; overall impressions; and preferences were all taken into account.²² Test users and takers were requested to suggest any changes to the items without

hesitation. After that, a discussion was held with the experts, and items were modified based on feedback from test users, test takers, and experts.²²

Stage II

The second pretest stage had a different group of five test takers. The questionnaire was pretested once again. As all the items were paradigmatic at this stage,²³ the parents were asked to rate the importance of the items on a five-point scale from (1) “not important” to (5) “extremely important.” The impact score was used to verify the face validity of all items.²⁴⁻²⁶ It was calculated using the proportion of test takers who rated it as important and the mean importance score given to the item. Items with an impact score >1.5 were retained. The same cognitive debriefing method as in the first pretest was repeated. The test takers’ feedback was used to modify the questionnaire appropriately. The trials helped us to understand the challenges that might arise while using the questionnaire.

Field Test

At this stage, 30 parents of first-time diagnosed children with ASD, who understood English, were requested to participate. If any child with ASD had major physical health issues, then his/her parent was excluded. The responses were used for item analysis. For this, Cronbach alpha was assessed, as the total number of response anchors in the present questionnaire was five.⁹ Item analysis was used for item reduction.

Results

Peer Review and Content Validation

Four items (items 12, 28, 35, and 37) found redundant were eliminated through the peer review process (Table S1). In the next two steps (CVR and CVI calculation), we used experts’ item ratings to decide about item inclusion/revision/exclusion. Table S2 reveals the CVR of the items as calculated in the first round. Thirty-four items with CVR <0.78 were removed in the first judgment stage. The remaining 34 of the 68 items had a CVR between 0.78 and 0.99, and a few were modified following expert advice and resubmitted

for the CVI calculation. The findings of the second stage of analysis (CVI) are in Table S3. It shows that 33 items had I-CVI values ≥ 0.78 , and one item that did not reach this threshold was deleted. Revisions to the items were made as and where suggested by the experts. The revised version of 33 items was again provided to the experts for the third stage of content validity. Each expert’s percentage of agreement at this point was >0.80 (Table S4); hence, all the experts were chosen for the next round of analysis. As seen in Table S5, in the final round of content validity, all the items except for three had an I-CVI of 0.889, or 1.00. The three items with CVIs <0.78 were eliminated. The CVI of the questionnaire by using the method of S-CVI/Ave, “summation of I-CVIs divided by the number of items,”¹⁵ (28.35/30) was 0.95, and by using the method of S-CVI/UA, “the proportion of items on an instrument that achieved a rating of 3 or 4 by all the content experts,” (15/30) was 0.50.¹⁵ The UA-CVI of 0.50 and the Ave-CVI of 0.95 indicate excellent content validity.^{19,27} The items with I-CVI values of 1.00 and 0.889 had ρ_c values of 0.002 and 0.018 and *K* values of 1.00 and 0.887, respectively. All items had *K* values >0.74 , which is considered excellent. The agreement on the comprehensiveness of the entire questionnaire was nine. The proportion of consensus regarding the comprehensiveness of the entire questionnaire was 1.00.^{15,27} After calculating ρ_c and *K*, three items were deleted from the questionnaire. They were “Behaves appropriately with the parent in family functions like birthday party,” “Comes to the parent when becomes upset,” and “Shows happiness with the parent.” Thirty items were thus ready to be pretested.

Stage I

Thirty items retained after CVI were administered by the target users (Table S6). Behavior coding revealed that 12 of the 30 items needed revision. Items were checked for ambiguities, directness of instruction, complexity of sentences, and understandability.

The following changes were made based on the suggestions of the test takers and test users and the consequent discussion with subject experts. The total number of items in the questionnaire

was specified in the introduction, and the term “item” was replaced with “statement” in the final questionnaire. Item 11 was deleted as it was similar to item 12. Item 2, “relates with the parent,” was modified to “relates his/her experience with the parent.” The phrase “day-to-day problems” was changed to “day-to-day concerns” in item 3. Items 4 and 19 were merged and modified to “greet while the parent leaves/returns home”. Item 4 was therefore removed. According to test takers, item 7 was problematic because picture sessions are done occasionally only in families. Hence, it was eliminated. In response to test takers’ suggestions, several changes were made to the word “play.” Item 8, “takes initiative to play with the parent”, was rephrased to “takes initiative to engage in some activities with the parent.” The phrase “moves away when the parent tries to play with him/her” was changed to “avoids engaging in parent-initiated activities” in item 10. Items 23 and 24 were removed as they were similar to item 22. The format was modified due to a scoring issue. To make the items more user-friendly, response anchors were written for each item with the respective numbers enclosed in brackets.

Stage II

The same process was repeated in the second stage. Additionally, the parents rated each item on a five-point scale to assess the questionnaire’s face validity. Table S7 displays the test takers’ ratings, which shows that the impact scores for all the items, except two, were >1.5 . These two items were thus excluded from further examination.

The final questionnaire was fine-tuned based on the expert panel’s advice. The phrase “through gesture” was removed from item 14. In item 20, examples that were in verb format were updated to noun format (feelings of anger, sadness, happiness, anxiety, etc.).

Overall, no difficulties were observed with the test administrators’ behavior, test environment, test fairness, test takers’ guessing, or their physical or mental conditions at any stage of the pretest. There were no culturally sensitive items. Several respondents declined to take the pretest because of language barriers. Test takers and users suggested

translating it into regional languages to benefit a larger target group, but due to time constraints, this could not be done. Although a few test takers stated that the questionnaire was lengthy, the majority stated that the items were highly relevant. The experts, however, judged the length of the questionnaire to be adequate. Twenty-three items were retained at this stage (Table S8). It took around 20–25 min to administer the questionnaire, and it captured the relevant issues for which it was developed.

Field Test

All 30 participants had an education level above the High School Certificate Examination (HSCE). The education level of five (16.66%) was between HSCE and below graduation, and 25 (83.33%) was above graduation. Ten (33.33%), 12 (40.00%), and eight (26.66%) were from rural, urban, and semi-urban locations, respectively. Ten (33.33%), 11 (36.66%), and nine (30.00%) belonged to low, middle, and high socioeconomic status, respectively.

Item analysis was carried out on the 23 items of the newly developed Nayak SER Questionnaire (NSERQ, available as an online-only supplementary file). Item 8 was reverse-scored. Cronbach's alpha revealed that the questionnaire had acceptable reliability, $\alpha = 0.931$. On examining the item-total correlations, retaining all the items appeared worthwhile. The Cronbach's alpha increased slightly by eliminating item 13, from 0.931 to 0.932, and the corrected item—total correlation of this item was 0.297. The corrected item—total correlation of the other items ranged from 0.436 to 0.871. Because this item's corrected item—total correlation score was quite close to the accepted range of $r = 0.30$ – 0.70 , it was not considered an outlier and was not eliminated. The mean of item 23 was 1.867, and the mean of the other items varied from 2.00 to 2.93. Though the acceptable range for the mean is >2 and <4 , the researchers did not consider it an outlier since its value was very close to the accepted range. We did not consider removing this item also because the Cronbach's alpha, if this item deleted, was lowered from 0.931 to 0.926. The skewness of all the items was <2 , and the kurtosis was <7 .

By keeping items 13 and 23, Cronbach's alpha indicated that the questionnaire had acceptable reliability, $\alpha = 0.931$, which is considered excellent for the clinical population. The total score would be determined by summing up all the responses to the questionnaire. The higher the score, the better the SER between the child and its parents.

Discussion

According to DSM-5,¹ there are two distinct domains of ASD. While the first domain which is three-dimensional pertains to social communication and social interaction, the second involves restricted and repetitive patterns of behavior, interest, or activities. NSERQ, however, assesses only one of the three dimensions of social communication and social interaction domain, and hence it cannot be used for diagnostic and/or severity assessment purposes. The scale can only be used by mental health professionals and researchers. The general population is cautioned against using the scale for self-diagnosis or other clinical purposes. Though this scale cannot be used as a stand-alone diagnostic tool, it can be used with other assessments to evaluate ASD, for research and therapeutic purposes only. The questionnaire's administration takes only around 20–25 min. The content validity process aided in determining the extent to which it captures the targeted construct, its relevant functions, and inferences that can be drawn from the resultant data.^{7,15} Content validity of the entire instrument (S-CVI) can be assessed by measuring both I-CVI and S-CVI using different approaches (e.g., CVR, CVI, S-CVI/Ave, S-CVI/UA). Because the scale-level CVI is an average score that might be skewed by outliers, the current research employed both techniques.

In the first judgment stage, a three-point scale was employed since a four-point scale results in a forced decision because there is no option to be unsure or neutral.¹⁵ Experts advised that 34 items be modified at this level. In the second stage, a four-point scale was adopted since it does not provide an ambiguous middle rating and also provides enough precise information to determine a valid CVI.¹⁶ I-CVI was checked for each item to determine

which item should be retained. Though 15 items underwent significant changes, those were not final at this stage. Another technique of Ave-CVI was calculated to determine if all of the experts should be chosen for the next round.¹⁵ They were all chosen for the next round since their proportion of agreement was >0.80 and supported that each item was crucial in evaluating the SER.

Various research has indicated a range of acceptable values for Ave-CVI ranging from 0.80 to 0.90.²⁷ S-CVI/UA estimated the overall content validity of the NSERQ to be 0.50. However, the S-CVI/Ave was 0.95. Although the Universal Agreement method considers only items with an I-CVI of 1.00 and may be considered more comprehensive than the Average approach, it may underestimate the overall questionnaire's content validity because the likelihood of achieving 100% agreement in all items decreases as the number of experts increases. The S-CVI/Ave approach, which is less constricted, may overestimate content validity since the numerator in the average technique is always greater than the Universal Agreement approach if all I-CVI values are not equal to 1.00. As a result, both the S-CVI/UA and S-CVI/Ave were computed; the NSERQ's overall content validity maybe somewhere in between.²⁸ The *K* result was interpreted following the recommendation given by Polit et al., specifically, "fair (*K* of 0.40–0.59), good (*K* of 0.60–0.74), or excellent (*K* > 0.74)."¹⁹ The proportion of consensus for comprehensiveness for the overall questionnaire was 1.00. Quantitative methods significantly supported individual items of the questionnaire, while qualitative methods such as cognitive interviews were employed to modify the questionnaire further.

There is no clear recommendation for the optimal number of revisions to ensure that a measure is well developed.²⁹ The final version of this questionnaire underwent two rounds of modification in the pretest stage. Such modifications addressed a wide range of issues, including the ease of comprehending and following instructions, challenges in taking the test, difficulties in the test-taking situation, and the ease of administration. Although the items were generic, the purpose was to specifically understand the SER of children on the autism spectrum.

A few limitations of this questionnaire are the small sample size in the pretest, its usability being restricted to those who understand English, the absence of an observation form, the likelihood of bias due to experts' subjective comments, and not evaluating the confirmatory factor analysis or construct validity due to time constraints.

Conclusion

The researchers developed the questionnaire through a three-round judgment-quantifying stage for content validation and two stages of pretesting. The I-CVI, S-CVI, ρ_c , K scores, and pretest results indicate that the current questionnaire containing 23 items is brief, clear, valid, and readily administrable for assessing the SER between children with ASD and their parents. This questionnaire demonstrated excellent content validity of individual items ranging from 0.89 to 1.00, and the entire questionnaire's content validity was S-CVI/UA = 0.50, S-CVI/Ave = 0.95. Assessment with this questionnaire would reveal SER dimensions that require greater attention, allowing mental health professionals to plan accordingly. This questionnaire has the potential to be utilized in both research and therapeutic settings, thereby facilitating mental health professionals to plan individualized interventions for children with ASD.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship and/or publication of this article.

Supplemental Material

Supplemental material for this editorial is available online.

ORCID iDs

Asmita Nayak  <https://orcid.org/0000-0002-8271-8213>

Rooplekha Khuntia  <https://orcid.org/0000-0002-1010-6728>

References

- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington, DC: APA, May 22, 2013.
- Dawson G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev Psychopathol* 2008; 20(3): 775–803.
- Rogers SJ and Vismara LA. Evidence-based comprehensive treatments for early autism. *J Clin Child Adolesc Psychol* 2008 Mar 3; 37(1): 8–38.
- Woods JJ and Wetherby AM. Early identification of and intervention for infants and toddlers who are at risk for autism spectrum disorder. *Lang Speech Hear Serv Sch* 2003 Jul 1; 34(3): 180–193.
- Juneja M, Mishra D, Russell PS, et al. INCLIN diagnostic tool for autism spectrum disorder (INDT-ASD): Development and validation. *Indian Pediatr* 2014 May; 51(5): 359–365.
- Mukherjee SB, Malhotra MK, Aneja S, et al. Diagnostic accuracy of Indian Scale for Assessment of Autism (ISAA) in children aged 2–9 years. *Indian Pediatr* 2015 Mar; 52(3): 212–216.
- Haynes SN, Richard D, and Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychol Assess* 1995 Sep; 7(3): 238.
- Berk RA. Importance of expert judgment in content-related validity evidence. *West J Nurs Res* 1990 Oct; 12(5): 659–671.
- Streiner DL, Norman GR, and Cairney J. *Health measurement scales: A practical guide to their development and use*. Oxford University Press, USA; 2015.
- Sireci SG and Faulkner-Bond M. Validity evidence based on test content. *Psicothema* 2014; 26(1): 100–107.
- Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Meas* 1995 Dec; 14(4): 5–8.
- Rubio DM, Berg-Weger M, Tebb SS, et al. Objectifying content validity: Conducting a content validity study in social work research. *Soc Work Res* 2003 Jun 1; 27(2): 94–104.
- Davis LL. Instrument review: Getting the most from a panel of experts. *Appl Nurs Res* 1992 Nov 1; 5(4): 194–197.
- Lawshe CH. A quantitative approach to content validity. *Pers Psychol* 1975 Dec 1; 28(4): 563–575.
- Almanasreh E, Moles R, and Chen TF. Evaluation of methods used for estimating content validity. *Res Social Adm Pharm* 2019 Feb 1; 15(2): 214–221.
- Lynn MR. Determination and quantification of content validity. *Nurs Res* 1986 Nov 1; 35(6): 382–386.
- McCain NL. A test of Cohen's developmental model for professional socialization with baccalaureate nursing students. *J Nurs Educ* 1985 May 1; 24(5): 180–186.
- Zamanzadeh V, Ghahramanian A, Rassouli M, et al. Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *J Caring Sci* 2015 May 31; 4(2): 165–78.
- Polit DF, Beck CT, and Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health* 2007 Aug; 30(4): 459–467.
- Beatty PC and Willis GB. Research synthesis: The practice of cognitive interviewing. *Public Opin Q* 2007 Jan 1; 71(2): 287–311.
- Presser S, Couper MP, Lessler JT, et al. Methods for testing and evaluating survey questions. *Public Opin Q* 2004; 68(1): 109–130.
- Cohen RJ and Swerdlik ME. *Psychological testing and assessment: An introduction to tests and measurement*. 6th ed. New York: McGraw-Hill, 2005.
- Presser S and Blair J. Survey pretesting: Do different methods produce different results?. *Sociol Methodol* 1994 Jan 1; 24: 73–104.
- Howard MC. Scale pretesting. *Pract Assess Res Evaluation* 2018; 23(1): 5.
- Juniper EF, Guyatt GH, Streiner DL, et al. Clinical impact versus factor analysis for quality of life questionnaire construction. *J Clin Epidemiol* 1997 Mar 1; 50(3): 233–238.
- Lacasse Y, Godbout C, and Series F. Health-related quality of life in obstructive sleep apnoea. *Eur Respir J* 2002 Mar 1; 19(3): 499–503.
- Polit DF and Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health* 2006 Oct; 29(5): 489–497.
- Rodrigues IB, Adachi JD, Beattie KA, et al. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. *BMC Musculoskelet Disord* 2017 Dec; 18(1): 1–9.
- Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. *Value Health* 2011 Dec 1; 14(8): 978–988.