

Assessing the Psychometric Properties of a Social Communication Test for Children with Autism Spectrum Disorder in Saudi Arabia

Samar Jaradat

*PhD, Department of Educational Studies, Univeristi Sains Malaysia, Malaysia
Jeddah, Saudi Arabia*

mssayaja@yahoo.com

ORCID: <https://orcid.org/0000-0001-7657-0382>

ABSTRACT

The proposed test is aimed to help in properly diagnosing children on the ASD spectrum and designing appropriate treatment plans for them. Due to the fact that social and communication skills are one of the major distinctive characteristics of ASD (National Research Council, 2001, DSM-5, 2013), this research study attempts to provide a test in Arabic language, the ASEST test, for the assessment of these skills. Adequate and accurate assessment of social and communication skills is considered to be one of the corner stone's in the assessment and diagnosis of ASD. The ASEST is designed to assess the social and communication skills of children with ASD in Saudi children between the ages of 6 and 11:11 years. It is critical to develop standardized tests in Arabic language that are linguistically and culturally appropriate. Due to the fact that social and communication skills are one of the major distinctive characteristics of ASD (National Research Council, 2001, DSM-5, 2013), this research attempts to provide a test in Arabic language for the assessment of these skills.

Keywords: Autism Spectrum Disorder; DSM-5; social communication skills; standardized test; assessment and diagnosis; psychometric properties

INTRODUCTION

To date, the diagnosis of Autism Spectrum Disorder (ASD) is still challenged with the huge dearth in the availability of standardized tests in Arabic language (Salhia, Al-Nasser, Taher, Al-Khathaami, & El-Metwally, 2014). This dilemma will continue to negatively affect the provision of appropriate services for individuals with ASD in the whole Arab region. Accordingly, it is critical to develop standardized tests in the Arabic language that are linguistically and culturally appropriate. Wetherby (2006) indicated that social-communication difficulties of children with ASD received remarkable attention in the past two decades. Therefore, these skills were included in the diagnostic criteria of ASD (American Psychiatric Association, 1994). The Arabic Social and Emotional Skills Test of Children with Autism Spectrum Disorder in Saudi Arabia (ASEST) focuses on the assess-

ment of the social and emotional skills of children on the spectrum. This test was developed and validated to assess the social communication skills of children with ASD in the city of Jeddah in Saudi Arabia who are between the ages of 6 years and 11 years 11 months old. It aims to help in properly diagnosing children on the spectrum and designing appropriate treatment plans for them. Gamliel and Yirmiya (2009) indicated that “the social behavior of children, adolescents, and adults with ASD differs qualitatively from that of their typically developing agemates” (as cited in Goldstein, Naglieri, & Ozonoff, 2009, p. 139). It is therefore of great importance to thoroughly assess these areas during the assessment of an ASD child. Gamliel and Yirmia (as cited in Goldstein, Naglieri, & Ozonff, 2009) indicated that children diagnosed with ASD exhibit severe difficulties in the areas of social and communication skills, unlike typically developing children. Since the impairment of social and communication skills

comprises one of the major defining criteria of ASD, all screening and assessment tests of ASD primarily focus on assessment of these skills (Goldstein, Naglieri, & Ozonff, 2009). The proposed Arabic test consists of four subtests, all designed to assess social communication skills.

PURPOSE OF RESEARCH

The purpose of this study is to develop and validate a test to assess the Social and Communication Skills of children with ASD in the city of Jeddah in Saudi Arabia who are between the ages of 6 years 0 months and 11 years 11 months old. The test is called “The Arabic Social Communication Skills for Children with Autism Spectrum Disorder Test” (ASEST). The researcher hypothesizes that children with ASD will manifest severely compromised social and communication skills than typically developing children. Reliability and validity of the test were examined. The proposed test is aimed to help in properly diagnosing children on the ASD spectrum and in designing appropriate treatment plans for them.

LITERATURE REVIEW

The literature review conducted provided information on the theoretical and conceptual framework of the study and the conceptual framework of this research. All theories pertinent to the development of social and behavioral skills were discussed. It also shed light on the available tests translated or validated in Arabic language that are used in the assessment process of children with ASD in Saudi Arabia. Moreover, it discussed the tests that are widely used in the assessment process of children with ASD overseas.

Past Research on Development of Tests to Assess Social Communication Skills of Children

Matson and Wilking (2009) indicated that many disorders are caused by deficits in social skills. These disorders include juvenile delinquency, ADHD, developmental disabilities, social and withdrawal, aggressive and antisocial behavior, mental health problems, challenging behaviors, and dropping out of school. Accordingly, the as-

essment and treatment of social deficits received significant interest and attention in clinical settings as well as in research (Matson & Wilking, 2009). This attention was triggered by Salter in 1949, Wolper in 1958 and 1969, and Zigler and colleagues in 1973 (Matson & Wilking, 2009).

Due to the pivotal role of social skills in the development of children from all ages, considerable amount of research has been published on tests that are designed for the assessment of these skills. Matson and Wilking (2009) conducted a thorough literature on 48 role-play and social skill tests. They focused on the assessment tests because “there have been few extensive reviews of trends in the field, particularly with respect to the assessment” (Matson & Wilking, 2009, p. 250). In their study, they reviewed the tests with regard to number of items included, items derivation, type of test format used to assess social skills, norming data, and psychometric data (reliability and validity). Out of the 48 tests reviewed, 40 tests met the criteria that these 40 tests are designed for assessment of social skills only. These tests were mostly developed in English language. However, there were a few tests from other countries and languages. These languages included the following: Japanese, Dutch, Hindi, Hebrew, Andalusian and Spanish, French, Turkish, Slovakian, Iranian, German and Portuguese. No tests were available in Arabic language.

With regard to tests’ norming Matson and Wilking (2009) reported that most of the tests reviewed were normed on children and adolescents. The ASEST test is designed for children and adolescents. As for the test formats, most of the social skills tests used forced-choice format. The same format is used for the ASEST test. The participants will be presented with different stimuli items and are requested to identify, recognize, analyze and express their opinions on the presented stimuli items.

Matson and Wilking further commented on the number of items in the tests reviewed and indicated that (typical test length was 25-75 items, which test developers note require a time of roughly 10-30 min to be completed (2009, p. 254). The initial draft of the ASEST tests consists of 68 test items and eight examples. The time expected to complete the test would be between 20 to 30 minutes.

The analysis of the psychometric data of the 48 tests reviewed by Matson and Wilking (2009) indicated that almost all tests used different reliability and validity methods. Reliability was tested using the following methods: interrater and test-retest. The validity was tested using construct validity and content validity methods. The ASEST will use similar reliability and validity methods.

Factor analysis was also done for 30 tests out of the 40 tests reviewed. The researcher will run factor analysis for the ASEST test too. According to Matson and Wilking (2009) “factor analysis is necessary for demonstrating construct validity and determining the internal structure of the scale with regard to the number and content of subscales” (p. 258).

The results pertinent to the raters and rating procedures in the 40 tests reviewed by Matson and Wilking (2009) indicated that most of the tests employed the following rating procedures: self-report, parent or caregiver report, and teacher report. For the ASEST test there will be one rater only, (the primary researcher).

Finally, Matson and Wilking (2009) stated that the development of social skills tests has received international attention and acceptance in the last 4 decades due to the significance of the “social skills as an important construct for children” (P. 257). This notion becomes more significant and vital in Arabic speaking countries due to the lack of availability of social skills tests. The ASEST is developed to help in filling this gap.

Matson and Wilking (2009) recommended developing tests for different purposes. The ASEST is developed for assessing social communication skills of children with ASD. This purpose is significant because the ASEST is not intended to be used as a general test with many other disorders. The ASEST will not only help diagnose children with ASD but also in designing appropriate intervention programs.

Conceptual Framework

The ASEST test is primarily constructed based on the theory of mind. In order to assess the social communication skills for children between the ages of 6 and 11 years 11 months, the ASEST

was conceptualized into the following four subtests respectively: Recognizing Facial Expressions and Emotions, Identifying and Describing Social and Communication Responses, Dealing with Social Situations, and Using Social Expressions. The following diagram demonstrates the framework design of the ASEST four subtests on the basis of the ToM:

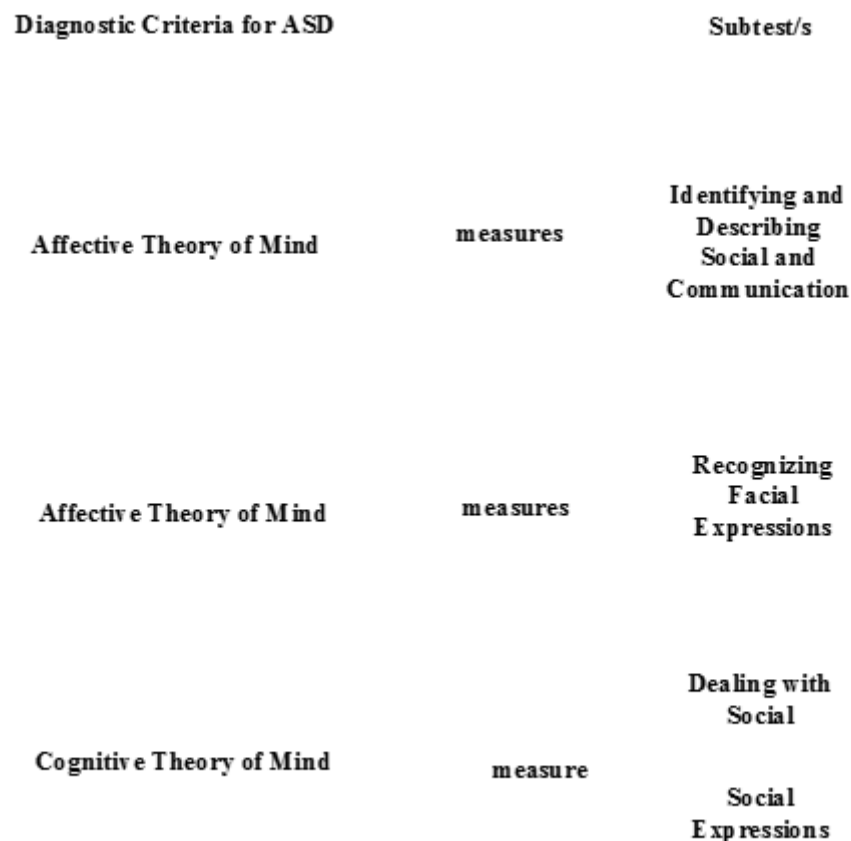


Figure 1: Conceptual framework of the ASEST test.

As shown in Figure 1 above, subtests one and two of the test are conceptually based on the affective ToM and subtests three and four of the test are conceptually based on the cognitive ToM.

Based on the above conceptual framework, the proposed test of the current study, (ASEST) is designed to assess the social communication skills of children with ASD in Saudi children between the ages of 6 and 11:11 years. The ASEST consists of the following four core subtests respectively: Recognizing Facial Expressions and Emotions, Identifying and Describing Social and Communication Responses, Dealing with Social Situations, and Using Social Expressions. The first subtest, Recognizing Facial Expressions and Emotions, evaluates the ability of the child to identify several facial expressions and feelings. The examiner tells the child that he will show her a picture depicting a specific feeling

for five seconds. The child is then presented with three pictures, two of the pictures are similar to the targeted feeling and one picture is different. The child is requested to point to the face that does not match the targeted feeling or emotion. The examiner says to the child “Which of these three pictures is different from the picture on the top?” This subtest is receptive in nature. The second subtest, Identifying and Describing Social and Communication Responses evaluates the receptive and expressive abilities of the child to identify the underlying cause of emotional reactions. The examiner points to the targeted picture and tell the child to point to the picture from a closed set of three pictures that makes the child feels that way. After the child points to picture, she is requested to name the targeted emotional reaction. The second subtest is both receptive and expressive in nature. The third subtest, Dealing with Social Situations, evaluates the ability

of the child to identify inappropriate social behaviors and then tells what should have been done to make the behavior socially appropriate. The examiner will show a picture to the child and asks the child "Is the person in this picture behaving appropriately?" If the child answers with Yes, the examiner has to move on to the next picture, but if the child answers with No the examiner says to the child "What is wrong and what should he/she have done?" The last subtest, Using Social Expressions, assesses the expressive ability of the child in different social contexts. The examiner will show the child a picture and then asks her "What are you supposed to say in this context or this situation?" This subtest is expressive in nature.

RESEARCH DESIGN

This study aimed to evaluate the social communication skills of Saudi Arabic-speaking children between the ages of 6:0 and 11:11 years. An Arabic test was developed for this purpose. A scale development and validation research using a cross-sectional research design which entails data collection at one point in time from a group of subjects selected based on demographic characteristics is employed. The cross-sectional assessment of social skills is conducted using the newly-developed "The Arabic Social Communication Skills for Children with Autism Spectrum Disorder Test" (ASEST). Typically developing children between the ages of 6:0 and 11:11 years and children diagnosed with ASD were recruited for the study from the public and private schools and special needs centers in the Northern Province in the city of Jeddah in Saudi Arabia.

Demographic Variables

There are five demographic variables in this study. The first variable was the developmental status of participants selected for the study that is either typically developing children or children with ASD. There were 120 normally developing boys and 120 normally developing girls between the ages of 6:0 and 11:11 years. Thirty-six children with ASD were also selected (three boys and three girls in each age group level). To design a valid instrument, it must reflect appropriate developmental levels of normally developing children. This information can be used later to establish basal and ceiling scales. The

second variable was the age of the participants. Children between the ages of 6:0 and 11:11 years were divided into the following age group levels: 6:0-6:11, 7:0-7:11, 8:0-8:11, 9:0-9:11, 10:0-10:11, and 11:0-11:11 years. This test was designed to measure the level of social communication skills acquired by participants. The ages 6:0-11:11 years are typical age range for early developing participants. The third variable is the gender of the participants. The fourth variable is the educational background of participants' parents or caregivers. This was done to evaluate the impact of the parents' educational level on the development of the children's social communication skills. The fifth is the academic performance of the participant.

Outcome Variables

The outcome variables included the overall performance of the participants on the AEST. Performance is defined here as the ability of the participant to correctly answer the test stimuli items of the four subtests. The test format used is the forced-choice format as according to a review by Matson and Wilkins (2009) on testing methods for children's social skills, the forced-choice has become by far the most popular test format given the trend over the last two decades. The second variable was the time duration the participant needed to complete the test.

Population

The study was conducted in the city of Jeddah in Saudi Arabia. The city of Jeddah was selected for the study because it is the second largest city in the kingdom and it represents the demographic structure of the Kingdom of Saudi Arabia. There are many Saudi families from all over Saudi Arabia who work and reside in Jeddah. Many families also bring their children with communication and behavioral problems for assessment and rehabilitation services because Jeddah has more special needs centers than other cities such as Al Taif, Tabuk, Dammam, and Arar. Finally, the city of Jeddah was chosen for the study because the researcher has more access to schools and special needs center in Jeddah than in any other city in Saudi Arabia.

This research will involve assessment of children with ASD as well as assessment of typically developing children in Jeddah. Two hundred and seventy-six children between the ages of 6:00 and 11:11 years will be recruited for the study.

Sampling and Participants

Typically-developing participants were randomly selected from the targeted schools. Every second student was chosen from a list of all typically developing students between the ages of 6:0 and 11:11 years. The inclusion criteria for participant selection included the following: all participants' parents were Saudis, participants are between the ages of 6:0 and 11:11 years, had normal cognitive ability and intact sensorineural abilities, were identified by the classroom teachers and parents as having average academic performance and reading skills, does not have any communication problems, and have age-appropriate social and behavioral development. All participants should have had no motor and sensory problems. The medical history of all participants was obtained from parents and school reports. Participant(s) with medical issues were excluded from the study as well as those who were on medication. Students who are reported by their teachers to have social communication problems were also excluded from the study. All the information above was collected from the students' records, teachers, and parents. Purposive sampling was conducted to select participants with ASD. The criteria for purposive sampling are: (a) Participants should be Saudi, (b) diagnosed with ASD, (c) between the ages of 6:00 and 11:11 years, and (d) enrolled in a treatment program. In this research, 276 students between the ages of 6:00 and 11:11 years will be recruited for the study. One hundred and twenty typically developing boys and one hundred twenty typically developing girls between the ages of 6:00 and 11:11 years. Thirty-six participants diagnosed with ASD will also be recruited for the study. The typically developing participants will be divided into the following age group levels: 6:0-6:11, 7:0-7:11, 8:0-8:11, 9:0-9:11, 10:0-10:11, and 11:0-11. Each age group level will consist of 20 boys and 20 girls (total 120 boys and 120 girls). As for the ASD group, each age group level will include three boys and three girls (total 36).

Psychometric Framework

The psychometric framework referred to in this research is the Standards for Educational and Psychological Testing by the American Education Research Association [AERA] (1999). The test development, validity, reliability, and test administration and scoring procedures chosen for this research refers to the Standards of AERA. Figure 2 below provides a summary of all the AERA standards employed for the ASEST test.

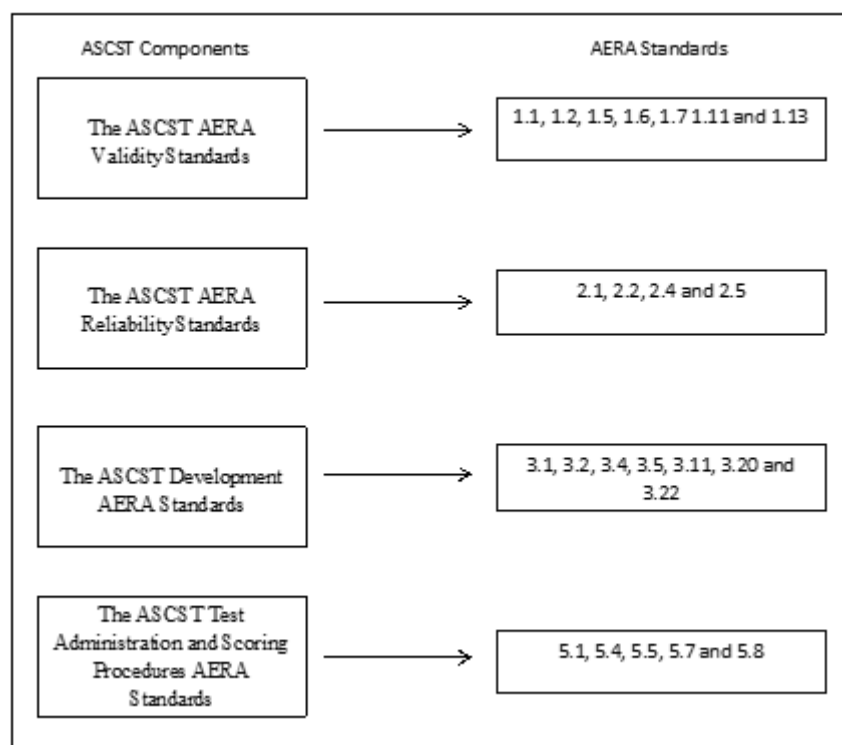


Figure 2 The AEST psychometric framework according to AERA

DISCUSSION

Normality Distribution of the AEST

The normality distribution of the AEST was done to examine the distribution of the items of the AEST. The skewness and kurtosis measures were tested, and the results of these measures revealed that data were normally distributed. The results further showed that the kurtosis values of the AEST are all falling in the kurtosis range. The analysis of skewness and kurtosis measures of the AEST showed that the content validity of the AEST was strong.

The normality curve of the AEST was examined for both groups, the typically-developing children and the ASD group. The results indicated that the typically-developing children scored significantly higher than their age-matched peers in the ASD groups.

The Evidence Pertinent to Content Validity of the AEST

Content validity in the current study is defined as the degree of representativeness of the content of test items to the behavior or characteristic to be measured (Crowl, 1996; Schiavetti & Metz, 1997).

The content validity of the AEST was tested and achieved through different methods including validation of the test items and picture items of the AEST, analysis of items difficulty and items discriminability, and the impact of age level on the performance of the typically-developing children on the AEST. Descriptive statistics was used to analyze the data.

The analysis of the results revealed that the content validity of the AEST was achieved. The initial draft of the AEST was carefully developed following the analysis of 20 validators (5 psychologists, 5 special education teachers, 5 parents and 5 speech-language pathologists). Five items were excluded from subtest one; six items were excluded from subtest two; two items were excluded from subtest three and no items were excluded from subtest four.

The exclusion of the items was done following the analysis of the validation results. The major criteria for the exclusion of the items were that these items were not age-appropriate, not linguistically-appropriate, not dialectically-appropriate and not applicable for the examinees.

As per the analysis of the picture validation results, one picture was removed from subtests one and two (receptive). No pictures were removed from subtests three and four. The pictures were removed because they do not meet the validation criteria.

The pictures were judged to be age-inappropriate, not clear, and culturally-inappropriate. These findings confirm the adequacy and appropriateness of the ASEST. The items and pictures agreed upon are all appropriate for testing what the test is designed for to test, i.e., the social communication skills of children with ASD between the ages of 6:00 and 11:11 years.

The results of the analysis of items discrimination and items difficulty revealed that the item discriminability distribution of the ASEST test was good and acceptable; item difficulty of the ASEST test was also judged to be balanced and acceptable. These findings again support the content validity of the ASEST.

The results of the descriptive statistics showed significant differences in the performance of the six cohort groups on the ASEST. These findings showed that the development of social communication skills of the Saudi children between the ages of 6:00 and 11:11 largely follows the same sequence of children in other languages such as English language (Tomasello, 1992; Nelson, 1985; Wetherby & Prizant, 2001).

The findings showed that children with ASD do exhibit significant deficits in the social communication profile. These deficits were highlighted by many theories and scholars for decades, such as theory of mind (Goldstein, Naglieri & Ozonoff, 2009), the joint attention theory (Trevvarthen, 1979), social-pragmatic view (Wetherby & Prizant, 2001), affective theory (Bauminger-Zviley, 2013); intersubjectivity theory (Trevvarthen & Atiken, 2001) and identification theory (Hobson & Meyer, 2005).

The analysis of the results further indicated that the performance of the Saudi children was consistent with the performance of their age-matched peers in English language when tested by tests designed for evaluating social communication skills such as the Social Responsiveness Scale (SRS) (Constantino & Gruber, 2005), and the Social Communication Questionnaire (SCQ) (Rutter, Bailey, Lord & Berument, 2003).

The analysis of the descriptive results of the ASD group yielded results obtained from children diagnosed with ASD in other languages in the same age group levels when evaluated for ASD such as the SEE test (Super Duper Publication, 2008).

The significant deficit in social communication skills is a major marker and criterion in the diagnosis of ASD as per the diagnostic criteria of DSM-5 (APA, 2013).

Another important evidence for the content validity of the ASEST comes from the available literature on social skills tests in English language and other international languages. This review is discussed in the following section.

The Evidence Pertinent to the Construct Validity of the ASEST

According to Graziano and Raulin (2007), construct validity is concerned with the theory or constructs the test results support. The ASEST test is designed to evaluate the social communication skills of children with ASD. The findings obtained from the current study revealed that the construct validity of the ASEST is high because these results showed that the test is valid, therefore the theories the test was based upon are then valid (Grazian & Raulin, 2007).

The construct validity reflects strongly the theories of the test, namely, theory of mind, joint attention and language (Bruner, 1983; Nelson, 1985; Tomasello, 1992). The ASEST was developed in line with the constructs proposed by these theories. The statistical results showed significant differences in the scores of the typically-developing children and the scores of children with ASD.

The findings further showed that ASEST was effective in detecting the group difference among

the six age groups targeted, therefore, the ASEST was representative of the domain being measured, the social communication domain.

These findings revealed that children with ASD exhibit significant deficits in their social communication profile. As such, these results support the assumption of the joint attention and language theory stated out by Wetherby and Prizant, (2001), where they indicated that children develop their linguistic symbols in line with the development of the social contexts with the adults. They further indicated that language acquisition and use is interrelated to the development of the social communication skills.

This explains the severe deficits of language problems and use in children with ASD. The theory of mind also postulated that children with ASD struggle remarkably in their linguistic abilities due to the severe deficits in the social communication skills. Goldstein, Naglieri and Ozonoff (2009) stated that the cognitive deficits affect the understanding of the social world. They further added that social communication deficits are evidently among the major and core difficulties in children with ASD.

The results of the current study fully support the hypothesis of the theory of mind. Additionally, the results of the current study are consistent with results of other tests in the field of special education in the United States, such as SEE (Super Duper Publication, 2008). All these findings showed that the construct validity of the ASEST was high and strong.

The severe deficits of the ASD participants in their communication skills in the present study might explain the association of joint attention with receptive language skills, expressive language skills, expressive vocabulary, development and use of grammar. Rollins and Snow (in press), Landry and Loveland (1988) and Loveland and Landry (1986) indicated that the development of receptive and expressive language skills is correlated with the development of proper attention skills (As cited in Wetherby & Prizant, 2001).

The SEE test was developed with the same construct of the ASEST that effective and adequate communication depend on effective and proper social communication abilities. The results of the ASEST markedly showed severe disruptions in the social communication skills of children with ASD. These disruptions explain that the

ASEST exhibit a high degree of construct validity in the same way as the construct validity of the SEE test.

The Evidence Pertinent to the Discriminant Validity of the ASEST

The performance of the typically-developing children was compared with that of the ASD children to test the discriminant validity of the ASEST. The Pearson Correlation was used to calculate means and standard deviation in the scores of the typically developing children group and the performance of the children with ASD group. The results showed significant differences between means and standard deviations of both groups in all subtests of the ASEST. The means of the typically developing children group was significantly higher than those for the group of children with ASD. The results further detected significant differences in the standard deviations of both groups. Evidently these results indicated significantly low correlation between both groups (typically developing children group and children with ASD group). These results suggested that the discriminant validity was achieved for the ASEST.

These findings were supported when the t test was used to see if the mean differences are significant or not. The t test showed that all the results were statistically significant between the typically-developing children group and the ASD children group across all the four subtests of the ASEST. These results were uncorrelated and thus indicated that the discriminant validity of the ASEST was high. McIntire and Miller, (2000) indicated that for discriminant validity to be achieved, scores should be uncorrelated.

The Evidence Pertinent to the Test-Retest Reliability of the ASEST

The fourth research question investigated the test-retest reliability of the ASEST when given to Saudi Arabic speaking children between the ages of 6:00 and 11:11 years. According to Schiavetti and Metz (1997) test-retest reliability is one of the major methods for testing reliability of tests.

Test-retest refers to the consistency and stability of results when the test is given more than one time (McLaughlin & Lewis, 2001; Graziano &

Raulin, 2007). The ASEST was administered twice over a one to two weeks interval time. The Pearson correlation was used to examine test-retest reliability of the ASEST. Thirty-six (18 boys and 18 girls) participants were randomly selected and were tested by the primary researcher. The ASEST was administered twice over a one to two weeks interval time.

The analysis of the results revealed that the test-retest reliability was achieved for all subtests and total score ($r = 0.852 - 0.994$). This result was statistically significant at $\alpha \leq 0.01$. This result is consistent with the reliability of other tests in the same domain such as the SEE where the test-retest reliability for the total raw score was .93, for the receptive raw score was .89, and for the expressive raw score was .88 (Super Duper Publications, 2008).

This result revealed that the test-retest was high for the ASEST and the results were stable over time and minimally affected by environmental or learning effects. Accordingly, the ASEST proved to be reliable in distinguishing children with ASD from those without ASD.

The Evidence Pertinent to the Interrater Reliability of the ASEST

Interrater reliability refers to the consistency of results of two raters or examiners where the consistency of both raters is measured. This means that the ratings or scores given by both raters of the same participant should be consistent (McLoughlin & Lewis, 2001; Graziano & Raulin, 2007). It is vital that both raters are blind to each other and do not know the ratings or scores of each other (Graziano & Raulin, 2007).

The results of the Pearson showed that the interrater reliability was achieved for all subtests and total score of the ASEST ($r = (0.893 - 0.968)$). This result was statistically significant at $\alpha \leq 0.01$. These results indicated a high level of correlation between the scores obtained by both the speech-language pathologist and the special education teacher. Overall, these findings are indicative of a high degree of scorer reliability of the ASEST.

These results were strongly consistent with the interrater results of the SEE test. In the SEE test, the interrater reliability for the receptive raw

score ranged between 0.99 to 1.00, for the expressive raw score, it ranged between 0.96 to 0.98 and for the total raw score, it ranged between 0.98 and 0.99.

The results are also in agreement with the interrater reliability of the Jordanian Expressive Vocabulary Test (JEVT) (Al-Sabi, 1997, in press). The JEVT is an Arabic test designed for the assessment of expressive vocabulary skills for children between the ages of 3:00 and 6:11 years. The interrater reliability of the JEVT ranged between 0.97 and 0.99. In summary, the ASEST is deemed appropriate and reliable in testing the social communication skills of ASD Saudi children between the ages of 6:00 and 11:11 years.

The Evidence Pertinent to the Split-Half Reliability of the ASEST

Although the split-half reliability method is strongly recommended and favored over other methods for testing reliability, it is not frequently used. The split-half reliability is used in testing the internal consistency of the reliability of the ASEST.

The Pearson correlation was used to test the split-half reliability. The correlations for all subtests were between ($r=0.800- 0.880$) and are all statistically significant at ($\alpha \leq 0.01$). This result confirms the reliability of the ASEST and supported the results obtained in the other reliability methods mentioned earlier, the test-retest and the interrater reliability. The summary of the three reliability methods used for examining the reliability of the ASEST revealed that the reliability of the ASEST was high and agrees with the standards of the AERA (AERA, 1999) as well as other tools in the same domain such as SEE (Super Duper Publications, 2008).

CONCLUSION

The main objectives of this research study have been achieved. A test has been developed, validated, and investigated. The major aim of the study was to examine all psychometric properties of the newly developed test, the ASEST. These properties included reliability, validity and internal consistency of the test. The results showed that ASEST exhibited a high degree of reliability and validity and can be used for the

evaluation of social communication skills of ASD Saudi children between the ages of 6:00 and 11:11 years. The findings of this study showed that the results were in full agreement with the proposed hypotheses of the investigator. The results revealed that ASD children present with severe deficits in the social communication skills. These results supported the diagnostic criteria of the DSM-5. Social communication deficits are a major landmark in the diagnostic criteria of the DSM-5 (APA, 2013). The results further supported the psychometric framework of the AERA. The AEST was based upon the standards of the AERA from the development stage to the validation stage. The findings further revealed that it is imperative to continue to develop and standardize assessment tools in Arabic language to improve and advance assessment and diagnostic services of children with ASD. The data collected from the study is also critical in establishing sound and effective intervention programs.

REFERENCES

- [1] Al-Sa'bi, Y. (2007). *The Jordanian Expressive Vocabulary Test Study*. Washington D.C: Howard University: unpublished dissertation.
- [2] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. American Educational Research Association.
- [3] American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. (4th ed.). Washington, DC: American Psychiatric Publishing.
- [4] American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- [5] Bauminger-Zviely, N. (2013). *Social and Academic abilities in children with high-functioning autism-spectrum disorders*. 72 Spring Street, New York, NY 10012: The Guilford Press.
- [6] Bruner, J. (1983). *Child's Talk: Learning to use Language*. New York: Norton.
- [7] Constantino JN, Gruber CP (2005). *Social Responsiveness Scale (SRS)*. Los Angeles: Western Psychological Services.
- [8] Crowl, T. (1996). *Fundamentals of educational research*. Madison, WI: Brown & Benchmark.
- [9] Gamliel, I., & Yirmaya, N. (2009). Assessment of social behavior in autism spectrum disorders. *Assessment of autism spectrum disorders*, 138-170.
- [10] Gilliam, J. E. (2001). *Gilliam Asperger's Disorder Scale: GADS*. Austin, TX: Pro-Ed.
- [11] Goldstein, S., Naglieri, J., & Ozonoff, S. (2009). *Assessment of autism spectrum disorders*. New York: Guilford Press.
- [12] Graziano, A., & Raulin, M. (2007). *Research methods*. Boston, Mass.: Pearson Allyn and Bacon.
- [13] Hobson, R., & Meyer, J. (2005). Foundations for self and other: a study in autism. *Developmental Science*, 8(6), 481-491. <http://dx.doi.org/10.1111/j.1467-7687.2005.00439.x>
- [14] Landry, S., & Loveland, K. (1988). Communication behaviors in autism and developmental language delay. *Journal of Child Psychology and Psychiatry*, 29(5), 621-634. <http://dx.doi.org/10.1111/j.1469-7610.1988.tb01884.x>
- [15] Lewy, A., & Dawson, G. (1992). Social stimulation and joint attention in young autistic children. *Journal of Abnormal Child Psychology*, 20(6), 555-566. <http://dx.doi.org/10.1007/bf00911240>
- [16] Loveland, K., & Landry, S. (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, 16(3), 335-349. <http://dx.doi.org/10.1007/bf01531663>
- [17] Matson, J. L., & Wilkins, J. (2009). Psychometric testing methods for children's social skills. *Research in Developmental Disabilities*, 30, 249-274.
- [18] McIntire, S., & Miller, L. (2000). *Foundations of psychological testing*. Boston, MA: McGraw-Hill.
- [19] McLoughlin, J. A., & Lewis, R. B. (2001). *Assessing students with special needs* (5th ed.). Upper Saddle River, N.J.: Merrill/Prentice Hall.
- [20] National Research Council. 2001. *Educating Children with Autism*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/10017>.

- [21] Rutter, M., Bailey, A., Lord, C., & Berument, S. K. (2003). *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.
- [22] Salhia, H. O., Al-Nasser, L. A., Taher, L. S., Al-Khathaami, A. M., & El-Metwally, A. A. (2014). Systemic review of the epidemiology of autism in Arab Gulf countries. *Neurosciences*, 19(4), 291–296.
- [23] Schiavetti, N., & Metz, D. (1997). Stuttering and the measurement of speech naturalness. In R. F. Curlee & G. M. Siegel (Eds.), *Nature and treatment of stuttering: New Directions* (2nd ed., pp. 398–412). Needham Heights, MA: Allyn & Bacon.
- [24] Tomasello, M. (1992). The social bases of language acquisition. *Social Development*, 1(1), 67–87. <http://dx.doi.org/10.1111/j.1467-9507.1992.tb00135.x>
- [25] Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 321–347). Cambridge: Cambridge University Press.
- [26] Trevarthen, C., & Aitken, K. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3–48. <http://dx.doi.org/10.1111/1469-7610.00701>
- [27] Wetherby, A. & Prizant, B. (2001). *Communication and symbolic behavior scales developmental profile- preliminary normed edition*. Baltimore, MD: Paul H. Brookes Publishing Co.
- [28] Wetherby, A., & Woods, J. (2006). Early social interaction project for children with autism spectrum disorders beginning in the second year of life: A preliminary study. *Topics in Early Childhood Special Education*, 26(2), 67–82.