

## **Analysis to Specificity and Sensitivity of Specifics Instruments Combined: Ados-2 and Adi-R, For Diagnostic Detection of Individuals with Autism Spectrum Disorders**

**PhD Manuel Ojea Rúa**  
University of Vigo  
Pontevedra, Spain

### **Abstract**

*Individuals with autism spectrum disorders (ASD) set up a very heterogeneous group of symptomatic criteria, which make up this disorder diagnosis, for this reason, it's necessary secure the effectiveness of multifunctional diagnosis through using of complementary instruments of reliability and internal consistency well contrasted. In this sense ADOS-2 and ADI-R tests shape two investigated instruments that shown an internal statistical goodness to facilitate precise diagnosis. This research contributes an analysis regarding the concordance of both measurement instruments to complementary use of both specific tests throughout this diagnostic process. The results found according statistical of Cohen's Kappa level analysis ( $k$ ) for  $N= 118$ , indicates that complementary agreement value between two tests is highly positive ( $k= .656$ ), of significant critical level:  $p= .00$ . These results show a high specificity and sensitivity of both tests combined use to facilitate the reliability and validity of diagnostic process.*

**Keywords:** Autism Spectrum Disorder; ADOS-2; ADI-R.

### **INTRODUCTION**

People with Autism Spectrum Disorder (ASD) make up a highly diagnostic criteria heterogeneous group, which, according to International Classification of Diagnostic and Statistical Manual of Mental Disorders (DSM-5 ®) of American Psychiatric Association (APA) (2013) and International Classification of Diseases (ICD-10) of World Health Organization (WHO, 1992), are synthesized regarding following specific diagnostic criteria: I) persistent deficit of reciprocal social and socio-emotional communication, II) nonverbal communication behavior deficits having along in social interaction, III) deficits into development, preserving and understanding of social relationships, and IV) deficits in restrictive and repetitive behaviors. Likewise, these criteria happen on different intensity frequencies levels, making up 3 levels or degrees of disorder severity, being 1: mildest level and 3: the most severe level, regarding diagnostic symptomatic groups of ASD.

This heterogeneous set of autism specific symptomatic characteristics needs the multifunctional employ of highly right instruments to make ASD early diagnosis with validity, consistence and reliability regarding ease the specific adapted psycho-educational intervention programs.

In this sense, Autism Diagnosis Observation Program test (ADOS-2) (Lord et al., 1999; Lord et al., 2012) and Autism Diagnosis Interview- Revised test (ADI- R) (Rutter et al., 2003) are the most important instruments to carry out diagnostic specific process to people with ASD.

Following study detailed analysis is completed.

### **ADOS-2 test**

Autism Diagnosis Observation Program, 2nd ed. (ADOS-2) (Lord et al., 1994) constitutes a standardized, semi-structured and just updated methodology to assess communication, social interaction, attention, imagination and

restricted and repetitive behaviors. This test, that been published by Western Psychological Services, concepts a subtle difference between autism spectrum disorder (ASD) and autism in strict sense throughout results measure process, differing to DSM-5 International Classification indications.

ADOS-2 test is made up of 4 basic modules:

- I) Module 1 is designed for 31- month- old children or also, for older children who use language communication phrases. It consists 10 activities formed by 34 elements belonging to effective-social components and restrictive and repetitive behaviors.
- II) Module 2 is destined for individuals who use language communication phrases. It comprises 14 activities with 29 differentiated codes for ASD specific diagnosis.
- III) Module 3 is intended for children with verbal fluidity and for adolescents, composing 14 activities with 29 measurement codes.
- IV) Module 4 is designed for adolescents and adults with verbal fluency, consisting 10-15 activities and 32 assessment codes structured severally to evaluate communication, reciprocal social interaction and repetitive and stereotyped behaviors.

ADOS-2 validation sample began with N= 98 participants, later expanded to N= 1.574 individuals and a replica made to N= 1.282 participants. Most the participants were autism diagnosis (45% -76%), following the autism spectrum diagnosis (7% -29%) and non-spectrum (17% -26%). The test validation on young children was developed to total of N= 182 participants with 360 evaluations, that is, most the children obtained 2 evaluations, which 25% of 182 participants were developing typical way, while 75% were risk of giving ASD diagnosis.

Internal consistency of Cronbach's Alpha for these tests in young individuals ranged between  $\alpha = .64$  and  $.88$ . Internal consistency for modules 1-3 obtained significantly high scores, located between  $\alpha = .88$ . and  $.90$ , compared to modulus 2 domain ( $\alpha = .88 - .90$ ). In module 4,  $\alpha = .75$  was obtained for communication:  $\alpha = .85$  for social interaction and  $\alpha = .47$  for repetitive and stereotyped behaviors.

Predictive sensitivity for test whole is between 60- 95%, with range laid between 75% and 100%, while for young children this sensitivity found is  $> 86\%$ . These data give the ADOS-2 test high reliability for diagnostic detection of ASD, especially owing its high sensitivity and concretion, however, the disorder diagnosis' multifunctional nature requires this test be complemented with other contrast instruments properly researched regarding to common concordance level.

### ***ADI-R test***

Diagnosis of Autism Revised Interview or ADI-R test (Rutter et al., 2003) makes up a comprehensive interview for teachers, families and caregivers, operationalized regarding the ASD diagnostic criteria, according to DSM-5 International Classification and the ICD-10- 10th revision (WHO, 1992), however, like the ADOS-2 test, conceptual evaluation process of this test differentiates between autism spectrum disorder (ASD 1-2), regarding autism itself (ASD- 3).

Test' initial empirical analyzes have been developed by Lord et al. (1994), who performed an experimental statistic comparison process on participants with ASD and other groups with intellectual and language disabilities, in which a total of 268 individuals participated along clinical sample for statistical validation, of which, 171 had ASD diagnosis. Results indicate that with well-adjusted cut-off points, scores offered a high level of reliability, located between 91%- 96% regarding diagnostic specificity, however, the getting of low sensitivities (44%-52%) stands out; however, the balance between specificity (60%- 62%) and sensitivity was significantly high: 77%-82%. Test' diagnostic specificity decreased when people were located on border of spectrum disorder (ASD- 1) (52%- 76%), but when limit of autism is used regarding severe symptomatic frequency (ASD- 3) this specificity was more adjusted: 81%- 83%. In synthesis, these results suggest that ADI-R test has 86% sensitivity and 92% high specificity.

Also, Vanegas et al. (2016) achieve a study of the ADI- R test specificity level. In their investigation, developed for N= 50, formed by ASD and typical groups, conclude the following resulting data for 3 domains of test: I) non-verbal communication, II) verbal communication, and III) repetitive and restrictive behaviors (SR) (see Table 1).

**Table 1: ADI- R specificity.**

Domains	ASD (n= 29)		TG (n= 21)		F	Partial $n^2$
	M	SD	M	SD		
Non-verbal communication	6.55	4.19	5.00	4.45	1.59	.03
Verbal communication	11.63	5.77	8.00	5.08	3.53 <sup>x</sup>	.10
Repetitive and restrictive behaviors	5.55	2.91	2.48	2.73	14.31**	.23

\*:  $p < .05$ , \*\*:  $p < .01$ , x:  $p < .10$

As can be seen, all test domains get high sensitivity of 78% and right specificity of 76.2%, with positive predictive values of 75% and negative predictive values of 80%. On the other hand, the data found by the ADI-R test keep stable, showing a high internal consistency of 67% of the boys and girls diagnosed with ASD. Stability increased to 78% in the repetitive and stereotyped behavior domain and to 88% when broader criteria were used in preschool-age participants (Soke et al., 2011).

These results have been later supported by different subsequent research studies (De Bildt et al., 2005; Falkmer et al., 2013), which verified empirical sensibility and specificity of ADI-R test, showing instrument validity and reliability for ASD specific diagnosis.

### **ADOS-2 and ADI-R tests**

ADI-R Diagnostic validity shown highly significant, but data improved when its statistical algorithms were examined relating to ADOS-2 test (Zander et al., 2015). Regarding this question, it's researched the specificity and sensitivity for both tests used complementary together.

Research sample for ADI-R \* ADOS-2 interaction, was initially analyzed for N= 268 participants from 18 to 47 age months, of which 171 presented ASD diagnosis. Study validity was determined for each element and domain that make up ADI-R test. Resulting algorithms showed significantly high specificities (91% -96%), but low sensitivities (44% -52%), however, when the diagnostic cut-offs were adjusted and were lower than cut-offs recommended, sensitivity balance enhanced very significantly (77% -82%), while it adjusted specificity (52% - 76%) for ASD intrinsic concept (ASD 1-2) of both tests. But, above all, when the more specific limit regarding Autism concept in strict sense (ASD- 3), sensitivity and specificity were best weighted: 81%- 94% to sensitivity and 81%- 83% to specificity.

Also, global data over all developmental subscales showed significant differences between ASD group and control group or typical group ( $p = .05$ ), but, when one data are combined between both tests, all levels height significantly about specificity and sensitivity. Thus, study interaction to both tests indicates following data: sensitivity= 77%- 80% for TEA concept (TEA 1-2), specificity= 85- 90% to TEA concept.

Comparative statistic Cohen's Kappa Coefficient ( $k = .23$  ( $p < .00$ ) for ASD 1-2 limit has been found and  $k = .30$  ( $p < .00$ ) when Autism (ASD- 3) was used in a specific conceptual sense.

Hence, the concordance level between adjusted ADI-R, combined with ADOS-2 for TEA concept, has been  $k = .31$  and  $.34$  respectively (to tests jointly:  $p < .00$ ). While each other to Autism (TEA-3) indicates agreement level:  $k = .15$ -  $.36$  (both,  $p \leq .04$ ).

Pearson correlation between both tests showed this same pattern of statistical data:  $r = .28$  ( $p = .00$ ) and  $r = .45$  ( $p < .00$ ).

Therefore, to deepen into aspects linked to the reliability and validity analysis of ASD diagnosis, in this research a study is carried out with the following general aims:

- 1) Analyze reliability and validity level to ASD diagnosis specific process.
- 2) Examine specificity and sensitivity level to ADOS-S and ADI- R tests each other for ASD diagnosis.

**METHOD**

**Research design**

Research design constitutes experimental analysis supported ASD specific diagnoses analysis corresponding to last 6 years, between 2015 and 2021. Resulting data have been analyzed throughout SPSS statistic.

**Participants**

IN this study a total of 118 specific diagnoses were collected. Recollecting data correspond to people aged between 0 and 23 years old, 41 participants aged 0-4 years, 56 aged 5-9 years, 11 aged 10-14 years and 10 greater 15 age years.

**Procedure and variables**

Specific diagnosis analysis (N= 118) subdivided into 3 categories, which correspond to 3 study' variables:

- 1) "No ASD": Negative diagnostic data regarding ASD.
- 2) "ASD": Positive diagnosis of ASD (ASD 1-2) according diagnostic tests analyzed.
- 3) "Autism": Positive diagnosis to Autism (ASD 3) according diagnostic tests concept used.

This differentiation between TEA and Autism is not compatible with DSM-5 Diseases International Classification, but it's reflected along evaluation criteria of analyzed tests: ADOS-2 and ADI-R.

Finally, concordance levels are observed, regarding specificity and sensitivity of both tests jointly.

**Data analysis**

SPSS statistical analysis is based on following data: 1) Frequencies contingencies analysis observed over each test: ADOS-2 and ADI-R relating to study variables: "No ASD", "ASD" and "Autism", and 2) Cohen's Kappa statistical comparative analysis to both test each other was performed.

**RESULTS**

Study resulting data are corresponded: 1) Study internal consistency, 2) Contingency' frequencies statistical analysis of both tests regarding participant age, and 3) Agreement comparative statistical level whole for ADOS-2 and ADI-R.

**Internal consistency**

Internal consistency analysis presents a Cronbach's Alpha shows significantly high consistency to whole study:  $\alpha = .903$ .

**Frequencies analysis**

Statistical frequencies data obtained for ADOS-2 variables: "No ASD", "ASD" and "Autism", regarding participants age can see in Table 2.

**Table 2: Frequencies value on ADOS-2 variables.**

		ADOS-2			Total
		No ASD	ASD (ASD1-2)	Autism (ASD 3)	
Age	0-4	22	15	4	41
	5-9	2	21	33	56
	10-14	4	7	0	11
	> 15	0	9	1	10
Total		28	52	38	118

As can be seen, ADOS-2 test evaluation data conclude that 28 participants were found on "No ASD" variable, of which, 22 participants have 0-4 years, 2 of 5-9 years, 4 of 10-14 years and 0 >15 years.

"ASD" variable was observed in 52 participants (15 aged 0-4 years, 21 of 5-9 years, 7 of 10-14 years and 9 >15 years).

“Autism” variable found in 28 diagnoses (4 of 0-4 years, 33 of 5-9 years, 0 of 10-14 years and 1 >15 years). Setting apart, ADI- R statistical frequencies results found over study variables regarding participant age can be seen in Table 3.

**Table 3: ADI- R frequencies.**

		ADI-R			Total
		No ASD	ASD (ASD 1-2)	Autism (ASD 3)	
Age	0-4	25	12	4	41
	5-9	5	35	16	56
	10-14	5	6	0	11
	> 15	0	9	1	10
Total		35	62	21	118

Indeed, regarding to ADI-R test variables, it's indicated a total of 35 participants were found as “No ASD” (25 of 0-4 years, 5 of 5-9 years, 5 of 10-14 years and 0 >15 years).

A total of 62 participants had "ASD" diagnosis variable, of which 12 participants had among 0-4 years, 35 of 5-9 years, 6 of 10-14 years and 9 >15 years).

Finally, it concludes to "Autism" diagnosis variable with total of 21 participants, 4 of 0-4 years, 16 of 5-9 years, 0 of 10-14 years and 1 >15 years.

### COMPARATIVE ANALYSIS

Tests variable data comparative analysis are developed.

Comparative statistical values, regarding Cohen's Kappa concordance indices analysis, it's possible deducing agreement level to diagnostic tests variables: "No ASD", "ASD" and "Autism" for both tests. Conclusive data indicate the following results to combined application of both tests (see Table 4).

**Table4: Cohen'sKappa.**

			ADIR-R			Total
			No ASD	ASD	Autism	
ADOS-2	No ASD	n	28	0	0	28
		%	23.7%	.0%	.0%	23.7%
	ASD	n	7	44	1	52
		%	5.9%	37.3%	.8%	44.1%
	Autism	n	0	18	20	38
		%	.0%	15.3%	16.9%	32.2%
Total		n	35	62	21	118
		%	29.7%	52.5%	17.8%	100.0%

Indeed, as can be seen in the cited Table, ADOS-2 and ADI-R tests agree in classifying “No ASD” diagnostic in 28 participants, “ASD” diagnostic in 44 and “Autism” diagnostic in 20 participants. Chi- Square value statistic of whole study:  $X^2 = 125.919$ ,  $df = 4$ ,  $p = .00$ . Pearson's correlation level gets:  $r = .827$ ,  $t = 15.822$ ,  $p = .00$  and, finally, Spearman correlation:  $s_p = .82$ ,  $t = 15.897$ ,  $p = .00$ .

In synthesis, agreement level between both instruments to diagnosis' variable type find significantly high critical levels. Therefore, it's conclude that both tests are highly complemented along diagnosis process, which greatly increases reliability and validity of diagnostic.

Hence, Cohen's Kappa index (k) can be observed in Table 5, indicates that concordance level is highly significant, showing high coherence between two tests.

**Table5: Cohen´s Kappa index.**

		Value	Error(a)	T(b)	Sig.
Concordance (ADOS-2-ADI-R)	Kappa	.656	.059	10.254	.000
Participants		118			

- a. Assuming alternative hypothesis.
- b. Using asymptotic standard error based over null hypothesis.

Comparative value level found is considered cohesion significantly:  $k = .656$ , with significant critical value (Sig= .00,  $T = 10.254$ ). It means the measures of both tests complement each other well-adapted way to determine the disorder diagnostic validity when used jointly. Thus, specificity and sensitivity rightly adjust to the reliability parameters disorder differential specific diagnosis.

This research corroborates that using the both specific tests: ADOS-2 and ADI-R complementary improve diagnostic consistence, owing highly descriptive and sensitive to ASD' specific criteria.

Then, as can be seen over Tables 4 and 5, data related with ADOS-2 test conclude that total of 28 participants (23.7%) were evaluated as "No ASD", 52 participants (44.6%) with "ASD" diagnostic, and 38 participants indicate like "Autism" (32.2%), while, ADI-R test detects 35 participants with "No ASD" diagnostic (29.6%), that's almost 6 points above the ADOS-2 test. Regarding variable "ASD", 62 participants are evaluated (52.5%), which involves almost 8 percentage points above the ADOS-2 test; and, regarding "Autism" diagnosis variable, 21 participants (17.8%) are concluded, that's, however, lower percentage than ADOS-2 test (almost 11.8 points below), which indicate that ADOS-2 test seems have higher level of specificity and sensitivity to most severe diagnosis: "Autism" variable that ADI-R test. However, but, both tests using complementary is essential to increase the validity and reliability of diagnostic process, as has been well corroborated by high concordance rates (k) of this study.

## DISCUSSION

There're also other specific diagnostic evaluation instruments as GARS-2-3 test (Gilliam Autism Rating Scale GARS-2, 2nd- 3rd ed.) (Gilliam, 2001; 2005; 2014). GARS test was high levels of reliability and validity, that has been analyzed along the work of 4 independent studies. Results found positive correlations with other ASD' diagnostic instruments including ADOS-2 test: 1) the Checklist of Autism Behaviors (3rd ed.) (ASIEP-3) (Krug, Arick y Almond, 2008), 2) the Carolina Autism Rating Scale (2nd ed.) (CARS-2) (Schopler et al., 2010), y 3) the Gilliam Asperger Syndrome Scale (GADS) (Gilliam, 2001). Total resulting mean scores found over correlations of the 4 tests indicate significantly high index ( $r^2 = .72$ ), range= .68- .77. Likewisemean autism index to 4 instruments is  $r^2 = .76$  and the range oscillates between .72- .83.

There're also other ASD' diagnostic evaluation instruments that can perform as complementary support along specific diagnostic processes. In this sense, Stewart and Lee (2017, p. 532) carry out an important review of ASD' diagnosis screening instruments used over different countries, checked from 1992 to 2015:

- ABC: Autism Behavior Checklist.
- 23Q: Twenty-Three Questions.
- M-CHAT: Modified Checklist for Autism in Toddlers.
- GARS: Gilliam Autism Rating Scale.
- SCQ: Social Communication Questionnaire.
- CARS: Childhood Autism Rating Scale.
- SCDC: Social Communication Disorder Checklist.
- AQ: Autism Spectrum Quotient.
- CABS-CV: Clancy Autism Behavior Scale.
- ASSQ-CV: Autism Spectrum Screening.
- SRS-CV: Social Responsiveness Scale.
- CAST: Childhood Autism Spectrum Test.

- CBCL: Child Behavior Checklist.
- ADEC-SP: Autism Detection in Early Childhood.
- ASQ: Autism Screening Questionnaire.
- SRS: Social Responsiveness Scale.

And, although there're variations over specificity sensitivity of confidence intervals regarding to different scales applied, this variation mainly owing to different adaptations and adjustments depending on needs of different countries where were implemented, as well as adjustments to specific contextual and individual characteristics, but all measurement instruments indicated can attend as complement to diagnosis specific essential tests currently contrasted: ADOS-2 and ADI-R tests.

## References

- American Psychiatric Association (2013). *Diagnostic and statistical Manual of mental disorders* (5<sup>a</sup> ed.). American Psychiatric Publishing. <https://www.psychiatry.org/psychiatrists/practice/dsm>
- De Bildt, A., Sytema, S., Kraijer, D., & Minderaa, R. (2005). Prevalence of pervasive developmental disorders in children and adolescents with mental retardation. *Journal of Child Psychology and Psychiatry*, 46, 275–286. <https://doi.org/10.1111/j.1469-7610.2004.00346.x>
- Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in Autism Spectrum Disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, 22, 329–340. DOI:10.1007/s00787-013-0375-0.
- Gilliam, J. E. (2001). *Gilliam Asperger's Disorder Scale*. Austin, TX: Pro-Ed. <https://www.txautism.net/evaluations/gilliam-aspergers-disorder-scale>
- Gilliam, J. E. (2005). Gilliam Autism Rating Scale GARS- 2 (2<sup>a</sup> ed.). México: PROED. [www.proedlatinoamerica.com](http://www.proedlatinoamerica.com)
- Gilliam, J. E. (2014). A test review. *Journal of Psychoeducational Assessment*, 35(3), 342-346. [sagepub.com/journalsPermissions.nav/journals.sagepub.com/home/jpa](http://sagepub.com/journalsPermissions.nav/journals.sagepub.com/home/jpa)
- Krug, D. A., Arick, J. R., & Almond, P. J. (2008). *Autism screening instrument for education planning* (3<sup>rd</sup> ed.). Austin, TX: Pro-Ed. <https://www.txautism.net/evaluations/autism-screening-instrument-for-educational-planning-third-edition-asiep-3-and-autism-behavior-checklist-abc>
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnosis Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders* 24, 659–685. DOI: 10.1007/BF02172145
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule: Manual*. Los Angeles, CA: Western Psychological Services. <https://link.springer.com/article/10.1023/A:1005592401947>
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2012) *Autism Diagnostic Observation Schedule (ADOS-2)*, 2<sup>a</sup> ed. Manual (Part I). Torrance, CA: Western Psychological Services. DOI:10.1177/0734282913490916
- World Health Organization (WHO) (1992). *International Classification of Diseases, 10<sup>th</sup> Revision (ICD-10)* World Health Organization. <https://www.cdc.gov/nchs/data/dvs/icd10fct.pdf>
- Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview Revised (ADI-R)*. Los Angeles, CA: Western Psychological Services. DOI: [https://doi.org/10.1007/978-1-4419-1698-3\\_894](https://doi.org/10.1007/978-1-4419-1698-3_894)
- Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Carolina Autism Rating Scale* (2<sup>a</sup> ed.). Los Angeles, CA: Western Psychological Services. DOI: 10.4236/eng.2011.34046
- Soke, G. N., Philofsky, A., Diguseppi, C., Lezotte, D., Rogers, S., & Hepburn, S. (2011). Longitudinal changes in scores on the Autism Diagnostic Interview–Revised (ADI-R) in pre-school children with autism. Implications for diagnostic classification and symptom stability. *Autism*, 15(5), 545-562. DOI: 10.1177/1362361309358332
- Stewart, L. A., & Lee, L. Ch. (2017). Screening for autism spectrum disorder in low- and middle- income countries: A systematic review. *Autism*, 21(5), 527-539. DOI: 10.1177/1362361316677025
- Vanegas, S. B., Magaña, S., Morales, M., & McNamara, E. (2016). Clinical validity of the ADI-R in a US based latino population. *Journal Autism Developmental Disorder*, 46, 1623- 1635. DOI 10.1007/s10803-015-2690-4
- Zander, E., Sturm, H., & Bölte, S. (2015). The added value of the combined use of the Autism Diagnostic Interview–Revised and the Autism Diagnostic Observation Schedule: Diagnostic validity in a clinical Swedish sample of toddlers and young preschoolers. *Autism*, 19(2), 187-199. DOI: 10.1177/1362361313516199