Reproducibility of an instrument for motor assessment of youth with autism

Reprodutibilidade de instrumento para avaliação motora de jovens com autismo

Carla Ferreira de Starmac*
Laura Bastianel
Taís Elena Heidrich
Cláudia Tarragô Candotti

Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

Date of first submission: March 26, 2023
Last received: December 23, 2023
Accepted: February 15, 2024

***Correspondence:** carla.starmac@yahoo.com.br

Abstract

Introduction: Individuals with autistic spectrum disorder (ASD) have motor impairments that precede communication and socialization disorders. Evaluative instruments compatible with the real possibilities and specificities of patients with ASD, and who quantitatively and qualitatively translate the data in which is wished to intervene with therapeutic actions, are important both in the scope of research and in the clinical evaluation of physiotherapists. **Objective:** To test the interobserver and intraobserver reproducibility of the instrument "Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder" (GMA-AUT checklist). **Methods:** The sample consisted of 34 individuals with ASD, aged between 6 and 18 years. The interobserver reproducibility was performed in a blinded manner by two physiotherapists experts in the ASD area of treatment. Intraobserver reproducibility was performed by one of the evaluators on two different days, with a gap of seven days and without access to data from the first evaluation. To verify the reproducibility, percentage of agreement and kappa statistics (k) were used, with the weighted kappa and, for the instrument scores, the intraclass correlation coefficient (ICC). **Results:** The GMA-AUT checklist showed excellent intraobserver agreement, with $k \geq 0.75$ and ICC > 0.75. Interobserver reproducibility ranged from good to sufficient agreement with k between 0.40 and 0.75 and ICC > 0.75 for the most part. **Conclusion:** The GMA-AUT checklist had excellent intraobserver reproducibility and, therefore, can be reliably used for assessments of individuals aged between 6 and 18 years with ASD.

Keywords: Autistic spectrum disorder. Physiotherapy. Reliability. Test reproducibility.

Resumo

Introdução: *Indivíduos com transtorno do espectro autista (TEA) apresentam comprometimentos motores que precedem os distúrbios de comunicação e socialização. Instrumentos avaliativos compatíveis com as reais possibilidades e especificidades dos pacientes com TEA, e que traduzam quantitativa e qualitativamente os dados nos quais se deseja intervir com ações terapêuticas, são importantes tanto no âmbito da pesquisa quanto na avaliação clínica do fisioterapeuta.* ***Objetivo:*** *Testar a reprodutibilidade interobservador e intraobservador do instrumento "Avaliação Motora Grossa de Crianças e Adolescentes com Transtorno do Espectro Autista" (checklist GMA-AUT).* **Métodos:** *A amostra foi composta por 34 indivíduos com TEA, com idade entre 6 e 18 anos. A reprodutibilidade interobservador foi realizada de forma cega por dois fisioterapeutas especialistas na área de tratamento do TEA. A reprodutibilidade intraobservador foi realizada por um dos avaliadores em dois dias distintos, com intervalo de sete dias e sem acesso aos dados da primeira avaliação. Para verificar a reprodutibilidade foram utilizadas a porcentagem de concordância e a estatística kappa (k), com kappa ponderado e, para os escores do instrumento, coeficiente de correlação intraclasse (CCI).* ***Resultados:*** *O checklist GMA-AUT apresentou excelente concordância intraobservador, com k ≥ 0,75 e ICC > 0,75. A reprodutibilidade interobservador variou de boa a suficiente concordância, com k entre 0,40 e 0,75 e ICC > 0,75 na maior parte.* ***Conclusão:*** *O checklist GMA-AUT apresentou excelente reprodutibilidade intraobservador e, portanto, pode ser utilizado de forma confiável para avaliações de indivíduos com idade entre 6 e 18 anos com TEA.*

Palavras-chave: *Transtorno do espectro autista. Fisioterapia. Confiabilidade. Reprodutibilidade dos testes.*

Introduction

Autistic spectrum disorder (ASD) is classified as a neurodevelopmental disorder.[1] It is characterized by behavioral and cognitive disorders that arise during the development period and involve significant difficulties in the acquisition and execution of intellectual, motor, language and social functions.[2] Its diagnosis is eminently clinical, based on criteria established by the Diagnostic and Statistical Manual of Mental Disorders (DSM 5), of the American Psychiatric Association, and by the International Classification of Diseases and Related Health Problems (ICD 11), of the World Health Organization (WHO).[3] Individuals with ASD present impairment in the motor aspect, more precisely in manual functions and global coordination, including laterality, changes in balance, posture and gait.[4] These impairments can persist throughout life and lead to motor deficits, such as altered gait, instability in postural control, balance problems, changes in tonus and deficiencies in gross motor coordination, such as independent walking and sitting.[5]

Motor development is not present in the analysis for the diagnosis of ASD because the motor assessment tests are long and demand a lot of time, making it difficult to apply in clinical dynamics.[6] Autonomous and functional movement is one of the great challenges for people with ASD; therefore, assessing, enhancing and incorporating it into daily life activities, based on a detailed analysis and intervention in the movement patterns presented since childhood, is a constant objective in physiotherapeutic practice, which is still recent.[7]

In a systematic review study, it was shown that standardized motor assessments used for patients with ASD allow identifying the presence of atypical motor development.[8] However, authors comment that the evaluations shown in literature have limitations, such as insufficient detail, demonstrating the lack of instruments for this specific public.[8] This fact was one of the main reasons for creating the instrument Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder (GMA-AUT checklist) (Appendix 1), which had its content validation previously tested.[9] However, for an instrument to be accepted in clinical practice and research, it must undergo rigorous statistical processes of validity and reliability, which demonstrate its measurement properties, with the analysis of reproducibility being a fundamental step.[10,11]

Therefore, since the GMA-AUT checklist has already been validated,[9] the objective of the present study was to test the interobserver and intraobserver reproducibility of the instrument. The hypothesis was that this instrument is reproducible, and that it can provide consistent responses over time when applied by the same evaluator or by different evaluators.

Methods

The research was characterized as a cross-sectional and prospective study with the aim of identifying whether the GMA-AUT checklist instrument would be reproducible to be used in the assessment of gross motor skills of autistic young people aged 6 to 18 years. This research was approved by the Ethics and Research Committee of the university where it was carried out, with assent number 5,485,604 (CAAE: 58836222.20000.5347), and followed the recommendations of Guidelines for Reliability Reports and Agreement Studies (GRRAS).[12]

The GMA-AUT checklist

This instrument consists of seventeen questions that are divided into static and dynamic assessment, and aims to evaluate the gross motor capacity of young people with ASD, providing quantitative and qualitative information.[9] Static assessment analyzes ground-sitting posture, balance and standing posture on different surfaces. The dynamic assessment includes postural changes, that is, it analyzes the movement from sitting to standing and vice versa; manual reception and bilateral ball kicking; going up and down ramps and stairs; and march. Each evaluated individual receives a score according to their performance in the "Do" or "Don't" sessions. If the individual performs the activity, the evaluator must also score the field "How to do it", reproducing this score in the "score" field (see Appendix 1). The instrument has the evaluation of 17 activities, each with sub-items that total 47 evaluated items.

Each of the evaluation response possibilities of the 17 activities of the GMA-AUT checklist, present in the three sections ("Do", "Don't" and "How to Do"), corresponds to the ordinal data provided by the instrument. The ability of the young person with ASD to carry out the assessed activity (Do/Don't) is observed based on the need or not to provide tips during the activity. The grading of these tips represents an ordinal scale: "No tip" = 5 points; "Verbal tip" = 4 points; "Gesture tip" = 3 points; "Modeling" = 2 points; "Partial physical tip" =1 point; "Does not" = 0 points. The motor performance (How Do You Do It) of young people with ASD is also assessed using a similar ordinal scale, which score can vary from 2 to 5 points, depending on the motor task. On these scales, the higher the score, the better the ability to perform and motor performance, respectively.

Furthermore, in addition to providing the qualitative characteristics of the assessment, the GMA-AUT checklist instrument generates quantitative information based on the scores present in the "Score Table", at the end of the instrument. Therefore, after carrying out all 17 activities and recording the points for the 47 items, the evaluator must calculate the final score (in percentage). The final GMA-AUT checklist score is the result of the sum of these points divided by 190 (which represents the instrument's maximum score).

The ideal of this evaluation is that the execution of the 17 activities is recorded on video for later analysis and scoring. The execution of these activities (motor tasks) does not always need to follow the same order during assessment, since all items are performed.[9]

Sample

The study population corresponded to young people with ASD, of both genders, aged between 6 and 18, enrolled in a special school located in Rio Grande do Sul (Brazil), which organizes classes in three school cycles: I (age 6 to 9), II (age 10 to 14) and III (age 15 to 18).

Sampling was intentional for convenience. Inclusion criteria were: individuals enrolled and attending school; aged from 6 years to 18 years and 11 months; and presenting a medical diagnosis (ICD F084) or have a hypothesis of diagnosis for ASD. Exclusion criteria were: being in a wheelchair; person's expression of unwillingness to take the test; and any situation that made it impossible to complete the assessment, such as complaints of discomfort, emotional situation on the day of test and any sensory alteration that modified the relational situation between evaluator/evaluated.

The sample size was defined considering the kappa statistic, with a two-tailed test, adopting a 90% power, assuming a null hypothesis with a kappa value = 0.00 and a detectable kappa of 0.60, consisting of a sample of at least 30 participants.[11]

Reproducibility and data collection

Data collection was carried out by one of the researchers (C.F.). For this collection, each of the young people with ASD, individually, carried out the complete assessment proposed by the GMA-AUT checklist. The entire assessment was filmed for later analysis purposes, according to the instructions of the instrument.[9]

Having possession of all the videos (one of each young person evaluated), two researchers (C.F. and T.H.), physiotherapists with experience in treating ASD, independently analyzed the videos using the GMA-AUT checklist. It is important to highlight that the videos were analyzed randomly by each of the researchers. There was an eight-hour training period for the evaluators that preceded the data collection period.

After finishing the video analysis, the following processes followed: analysis of interobserver reproducibility: it consisted of analyzing the agreement of the results of the evaluations with the GMA-AUT checklist between researchers (C.F. and T.H.); and analysis of intraobserver reproducibility: it consisted of analyzing the agreement of the results of two assessments with the GMA-AUT checklist carried out by researcher C.F, with the second assessment occurring seven days after the first. It should be noted that at the time of the second evaluation, researcher C.F. did not have access to the previous evaluation. The time interval (7 days) was chosen so that there would be no memory of the answers from the first day, and in accordance with Sim and Wright.[11]

Data analysis

Data were analyzed using SPSS software (Statistical Package for Social Sciences – version 26.0). To characterize the sample, a frequency table was used. The null hypothesis was that there would be no agreement [kapa = 0 and intraclass correlation coefficient (ICC) = 0] between the two evaluation situations (intraobserver and interobserver analyses). The level of statistical significance for all analyzes was $\alpha < 0.05$.

The weighted kappa and the percentage of agreement (%C) were used to evaluate the agreement between the inter and intraobserver reproducibility analyzes[12] in relation to ordinal data generated by the GMA-AUT checklist. For weighted kappa test interpretation, was verified that scores greater than or equal to 0.75 would represent excellent agreement; scores between 0.40 and 0.75, good or sufficient agreement; and scores lower than 0.40, poor agreement.[13] To consider the instrument reproducible, the minimum criteria of 0.50 for percentage of agreement and 0.40 for kappa were adopted.[14]

The ICC, with the standard error measurement (SEM) and the minimum detectable error (MDC – minimum detectable change) were used to evaluate the reproducibility in relation to the total score for each of the 17 activities, as well as the final score of the instrument, both in inter and intraobserver analysis.[15,16] The ICC score was classified as: poor (ICC < 0.40), moderate (ICC between 0.40 and 0.75) or excellent (ICC > 0.75).[16,17] SEM was estimated using the formula: $SEM = SD \sqrt{(1-ICC)}$, where SD is measurement standard deviation. MDC was estimated based on a 95% confidence interval, where $MDC = 1.96 \times SEM$.[17] For the instrument scores to be reproducible, we considered an ICC of at least 0.40.[18]

Results

Thirty-four individuals with ASD participated in this study (Table 1). There was no sample loss. The evaluations were carried out by two evaluators.

Table 1 - Sample characterization (n = 34)

Characteristic	n (%)
I Cycle (age 6 to 9 years)	15 (44.11)
II Cycle (age 10 to 14 years old)	15 (44.11)
III Cycle (age 15 to 18 years old)	4 (11.78)
Sex	
Male	25 (73.53)
Female	9 (26.47)
Comorbidity	
Yes	22 (64.71)
No	12 (35.29)
Use of medication	
Yes	27 (79.41)
No	7 (20.59)
Physiotherapy	
Yes	33 (97.05)
No	1 (2.95)
Extra class physical activities	
Yes	33 (97.05)
No	1 (2.95)
Other therapies	
Yes	21 (61.77)
No	13 (38.23)
Verbal language	
Absent	20 (58.83)
Present	14 (41.17)

In the sample, male subjects aged less than 14 years prevailed. Among comorbidities, there was a predominance of epilepsy. The most used medication was Respiridona. Most individuals in the sample do not take any complementary therapy; among those who take, the predominant therapy was occupational therapy.

Regarding the execution of GMA-AUT checklist activities, young people were evaluated with "Does" and "Does not". From 17 proposed actions by the instrument, all individuals in the Cycle III (n = 4) fully completed the assessment (100%); in Cycle II (n = 15), two individuals (13.33%) did not perform any of the actions, one did not perform Q2, Q10, and Q13, and two did not perform Q3. In the Cycle I (n = 15), seven individuals (46.66%) were evaluated with "does not perform"; with only one individual in Q2, Q3 and Q13; three individuals in Q9 and four individuals in Q10.

Regarding intraobserver reproducibility (Table 2), the GMA-AUT checklist obtained an agreement percentage of 100% in 29 questions; 17 were above 91.2%, and in only one question the result was 88.2%. As for the weighted Kappa, the dominant values were very close to 1, with the lowest being 0.80, representing excellent intraobserver reproducibility.[13] For total sums of each evaluated question of the GMA-AUT checklist (Table 3), ICC presented values > 0.75, demonstrating an excellent intraobserver reproducibility in the final instrument scores.

Table 2 - Intraobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder checklist (n = 34)

Question	Weighted kappa	95% Confidence interval	% of agreement	p-value
1	1.00	1.00 – 1.00	100	<0.01
1a	1.00	1.00 – 1.00	100	<0.01
1b	1.00	1.00 – 1.00	100	<0.01
1c	0.92	0.78 – 1.06	97.1	<0.01
2	0.98	0.95 – 1.01	97.1	<0.01
2a	0.95	0.86 – 1.04	97.1	<0.01
3	0.98	0.95 – 0.01	97.1	<0.01
3a	1.00	1.00 – 1.00	100	<0.01
4	1.00	1.00 – 1.00	100	<0.01
4a	1.00	1.00 – 1.00	100	<0.01
5	1.00	1.00 – 1.00	100	<0.01
5a	0.86	0.63 – 1.08	97.1	<0.01
6	1.00	1.00 – 1.00	100	<0.01
6a	1.00	1.00 – 1.00	100	<0.01
7	1.00	1.00 – 1.00	100	<0.01
7a	1.00	1.00 – 1.00	100	<0.01
8	1.00	1.00 – 1.00	100	<0.01
8a	0.87	0.73 – 1.02	94.1	<0.01
9	1.00	1.00 – 1.00	100	<0.01
9a	1.00	1.00 – 1.00	100	<0.01
10	0.98	0.95 – 1.01	97.1	<0.01
10a	1.00	1.00 – 1.00	100	<0.01
11	1.00	1.00 – 1.00	97.1	<0.01
11a	1.00	1.00 – 1.00	100	<0.01
11b	0.87	0.71 – 1.02	100	<0.01
11c	0.91	0.80 – 1.03	94.1	<0.01
11d	1.00	1.00 – 1.00	100	<0.01

Table 2 - Intraobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder checklist (n = 34) (continued)

Question	Weighted kappa	95% Confidence interval	% of agreement	p-value
11e	0.81	0.61 – 1.01	91.2	<0.01
11f	1.00	1.00 – 1.00	100	<0.01
12	0.95	0.86 – 1.04	97.1	<0.01
12a	1.00	1.00 – 1.00	100	<0.01
13	1.00	1.00 – 1.00	100	<0.01
13a	1.00	1.00 – 1.00	100	<0.01
14	0.89	0.74 – 1.04	94.1	<0.01
14a	0.96	0.90 – 1.03	97.1	<0.01
14b	0.96	0.88 – 1.03	97.1	<0.01
15	1.00	1.00 – 1.00	100	<0.01
15a	1.00	1.00 – 1.00	100	<0.01
15b	0.92	0.77 – 1.06	97.1	<0.01
16	1.00	1.00 – 1.00	100	<0.01
16a	1.00	1.00 – 1.00	100	<0.01
16b	0.93	0.80 – 1.06	97.1	<0.01
16c	1.00	1.00 – 1.00	100	<0.01
17	1.00	1.00 – 1.00	100	<0.01
17a	1.00	1.00 – 1.00	100	<0.01
17b	0.92	0.78 – 1.07	97.1	<0.01
17c	0.80	0.62 – 0.98	88.2	<0.01

Table 3 - Intraobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder scores (n = 34)

Item	Intraclass correlation coefficient	Standard deviation	Standard error of the mean	Minimum detectable change
Q1	0.98	0.71	0.08	0.23
Q2	0.99	2.08	0.17	0.48
Q3	0.99	2.15	0.09	0.26
Q4	1.00	1.09	0.00	0.00
Q5	0.98	1.35	0.17	0.47
Q6	1.00	1.51	0.00	0.00
Q7	1.00	1.25	0.00	0.00
Q8	0.99	1.59	0.12	0.34
Q9	1.00	2.72	0.00	0.00
Q10	0.99	1.59	0.10	0.28
Q11	0.99	2.13	0.21	0.59
Q12	0.98	1.46	0.17	0.48
Q13	1.00	2.68	0.00	0.00
Q14	0.96	2.30	0.44	1.24
Q15	0.99	2.67	0.16	0.46
Q16	0.99	1.89	0.08	0.23
Q17	0.98	1.55	0.22	0.61
Total	0.99	16.33	0.89	2.48

The GMA-AUT checklist obtained in the interobserver evaluation (Table 4) an agreement percentage above 50% in the majority (43 questions), with only four results below this value (in questions 1, 2, 7a, 17). As for the weighted Kappa, 21 questions showed excellent reproducibility; 22 questions presented values corresponding to good or sufficient reproducibility; and four questions (1, 7a, 16c, 17c) had values considered as poor reproducibility.

For the ICC (Table 5), most ICC values were > 0.75, establishing an excellent ICC, except for question 1, which was < 0.40 (weak ICC).

Table 4 - Interobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder checklist (n = 34)

Question	Weighted kappa	95% Confidence interval	% of agreement	p-value
1	0.13	-0.14 – 0.40	17.6	0.63
1a	0.79	0.53 – 1.06	94.1	<0.01
1b	0.63	0.35 – 0.91	85.3	<0.01
1c	0.78	0.54 – 1.06	91.2	<0.01
2	0.51	0.31 – 0.70	47.1	<0.01
2a	0.58	0.40 – 0.75	67.6	<0.01
3	0.44	0.19 – 0.68	50.0	<0.01
3a	0.83	0.67 – 0.99	91.2	<0.01
4	0.60	0.34 – 0.86	79.4	<0.01
4a	0.67	0.42 – 0.92	85.3	<0.01
5	0.83	0.67 – 0.99	88.2	<0.01
5a	0.75	0.47 – 1.04	91.2	<0.01
6	0.81	0.68 – 0.94	79.4	<0.01
6a	0.43	0.20 – 0.66	64.7	<0.01
7	0.87	0.74 – 1.00	88.2	<0.01
7a	0.16	0.15 – 0.31	41.2	0.02
8	0.55	0.33 – 0.78	79.4	<0.01
8a	0.66	0.50 – 0.82	82.4	<0.01
9	0.89	0.79 – 1.00	88.2	<0.01
9a	0.93	0.83 – 1.03	94.1	<0.01
10	0.95	0.89 – 1.01	94.1	<0.01
10a	0.93	0.84 – 1.02	94.1	<0.01
11	1.00	1.00 – 1.00	100.0	<0.01
11a	0.66	0.47 – 1.27	97.1	<0.01
11b	0.68	0.41 – 0.94	85.3	<0.01
11c	0.51	0.26 – 0.76	67.6	<0.01
11d	0.62	0.42 – 0.83	73.5	<0.01
11e	0.52	-0.28 – 0.38	55.9	0.76
11f	0.87	0.74 – 1.00	91.2	<0.01
12	0.72	0.52 – 0.92	79.4	<0.01
12a	0.55	-0.50 – 1.10	94.1	<0.01
13	0.74	0.62 – 0.86	58.8	<0.01
13a	0.85	0.76 – 0.93	79.4	<0.01
14	0.89	0.74 – 1.09	94.1	<0.01
14a	0.84	0.70 – 0.97	85.3	<0.01

Table 4 - Interobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder checklist (n = 34) (continued)

Question	Weighted kappa	95% Confidence interval	% of agreement	p-value
14b	0.44	0.20 – 0.68	61.8	<0.01
15	0.84	0.70 – 0.97	85.3	<0.01
15a	0.96	0.90 – 1.03	97.1	<0.01
15b	0.57	0.36 – 0.78	67.6	<0.01
16	0.85	0.71 – 0.99	85.3	<0.01
16a	0.74	0.58 – 0.90	76.5	<0.01
16b	0.43	0.21 – 0.64	67.6	<0.01
16c	0.17	-0.35 – 0.38	47.1	1.50
17	0.79	0.60 – 0.99	91.2	<0.01
17a	0.86	0.71 – 1.02	91.2	<0.01
17b	0.41	0.20 – 0.63	61.8	<0.01
17c	0.16	-0.11 – 0.44	52.9	0.19

Table 5 - Interobserver reproducibility of the Gross Motor Assessment of Children and Adolescents with Autism Spectrum Disorder scores (n = 34)

Item	Intraclass correlation coefficient	Standard deviation	Standard error of the mean	Minimum detectable change
Q1	0.31	2.89	2.40	6.65
Q2	0.81	4.29	1.85	5.15
Q3	0.81	3.65	1.58	4.38
Q4	0.72	3.25	1.72	4.77
Q5	0.93	3.23	0.83	2.32
Q6	0.83	3.46	1.40	3.88
Q7	0.81	3.44	1.49	4.13
Q8	0.83	3.21	1.29	3.59
Q9	0.99	3.62	0.32	0.89
Q10	0.99	3.93	0.32	0.91
Q11	0.70	6.11	3.30	9.14
Q12	0.88	2.73	0.92	2.55
Q13	0.96	3.93	0.72	2.00
Q14	0.94	3.58	0.85	2.37
Q15	0.98	3.82	0.52	1.46
Q16	0.87	3.44	1.23	3.41
Q17	0.87	3.73	1.32	3.68
Total	0.96	17.35	3.10	8.60

Discussion

In this study, the reproducibility of a gross motor assessment instrument for young autistic people, the GMA-AUT checklist, was analyzed. The instrument demonstrated excellent intraobserver reproducibility, with high %C, weighted kappa and ICC values (Tables 2 and 3), providing valuable material for physiotherapeutic assessment. In the interobserver evaluation, the result was between good and sufficient reproducibility (Tables 3-5). Questions with weighted kappa below 0.40 (poor reproducibility) were 1, 7a, 16c and 17c; and questions

that showed a low percentage of agreement were 1, 2, 7a and 16c. Questions 1 and 2 indicate whether or not the child does the motor skill remaining seated on the ground and standing on a stable surface, respectively. Questions 7a, 16c and 17c punctuate the motor patterns used for performing actions: move from standing to sitting position; go up and down the ramp with or without postural compensations.

Instruments that assess postures and movements (static and dynamic assessment) are subject to evaluators' interpretation subjectivity, especially when they involve the assessment process dynamics.[19] The results of such assessments are also influenced by fluctuating motor patterns in autistic children and adolescents. Even if the instrument indicates scoring the best performance,[9] some patterns may go unnoticed if they appear subtly, challenging interrater observational interpretations.[20] A study that evaluated the reliability of several instruments adds that, when tested, verbal cues, pauses, acclimatization, cards and even the environment itself were considered in the tests; however, there is little detail about how these variables were included in their research and to what extent they influence the evaluation process and instrument results.[20]

In a survey that verified the intraobserver and interobserver reliability for the Test of Gross Motor Development 3 (TGM-3), which includes a visual support protocol to facilitate task comprehension for children aged 3 to 10 years with autism, the conclusion was that most ICCs for intraexaminer reliability were higher than interexaminer reliability.[19] This finding corroborates previous evidence that an examiner is more likely to have higher scores for the intrarater reliability test, as opposed to the interrater reliability test, where differences in individual viewpoints and methods between raters can result in lower scores,[19] and wich also occurs in the validation of instruments from others areas, such as postural assessment, for example.[21]

A systematic review study that evaluated which instruments were most used in autistic children and adolescents and which accommodations were necessary to complete the tests[20] showed that the Test of Gross Motor Development 2 (TGMD2) and the Movement Assessment Battery for Children (MABC) were the most frequently reported, followed by the Bruininks–Oseretsky Motor Proficiency Test (BOT) and the Peabody Developmental Motor Scales-2 (PDMS-2). On the other hand, the literature also points out that the most used instruments to assess the public with ASD are not specific for this disorder, and may underestimate or overestimate the analyzed variables. In addition, the only instruments for individuals with ASD tested for validity and reliability were the BOT-2 and the Vineland Adaptive Behavior Scales. Second Edition (Vineland-II), which were validated for ASD populations.[22,23]

Regarding the performance of the sample in questions under the aspect "Does" or "Does not", children from Cycle I presented greater difficulty in tasks that required balance such as stand on a stable and semi-unstable surface, kick the ball with your right and left foot and jump, being marked "Does not" in questions 2, 3, 9, 10 and 13 for seven children. In Cycle II, only two children did not perform some of the actions proposed by the instrument as in the previous cycle. In Cycle III all children completed the evaluation. These results corroborate the literature, which shows that there is a linear trend towards an improvement in balance from 2 to 12 years of age.[24,25] Despite the literature showing that during school age autistic children have several difficulties with gross motor skills involving body balance, such as running, jumping, hopping,[26] in the present study the GMA-AUT showed that children with ASD, when compared to each other, demonstrated to have the same chronological line of development of motor skills according to neurological maturation and motor experiences.

The main limitation of this study is associated with both the subjective characteristic of the motor assessment proposed by GMA-AUT checklist, and the fluctuation of motor patterns in autistic children and adolescents. The lack of a user manual was a downside once it could guide evaluators in both the collection and analysis of videos. Since tips and the environment are identified as a factor that influences the evaluation,[20] it is believed that a manual of guidelines for the GMA-AUT checklist would have equalized the evaluators' doubts. In order to minimize interpretation problems, in this research the evaluators received eight hours of prior training. However, a positive feature of the GMA-AUT checklist is the possibility of providing qualitative information about how the individual performs the motor tasks, and not just inform whether he or she was able to perform it or not.[9]

Thus, considering that the GMA-AUT checklist is an instrument created for individuals with autism, that has validated content and with excellent intra-observer

reliability indexes and good or sufficient for inter-observer reliability, it constitutes a tool for clinical practice and can be used safely by professionals who care for autistic patients. In the academic field, the GMA-AUT checklist can be useful in research that seeks to identify the results of the treatment of patients with autism, or in studies tracking the motor patterns of individuals with autism.

Conclusion

Excellent levels of intraobserver reliability were obtained for the GMA-AUT checklist. For interobserver reproducibility, the instrument had good or sufficient reproducibility. The lowest ICCs were only in four questions out of the 47 that the instrument presents (17 questions and 30 subitems). These results demonstrate the reliability of the GMA-AUT checklist instrument, which constitutes an excellent tool for the clinical assessment of young autistic people.

Authors' contributions

LB, TEH and CTC were responsible for the study conceptualization. CFS and TEH, for data collection training, data collection and analysis, while LB supervised the analysis. CFS, LB and CTC wrote and reviewed the manuscript. All authors approved the final version.

References

1. Lum JAG, Shandley K, Albein-Urios N, Kirkovski M, Papadopoulos N, Wilson RB, et al. Meta-analysis reveals gait anomalies in autism. Autism Res. 2021;14(4):733-47. DOI

2. International Classification of Diseases, Eleventh Revision (ICD-11). Geneva: World Health Organization; 2022. Full text link

3. Morocho Fajardo KA, Sánchez Álvarez D, Patiño Zambrano VP. Perfil epidemiológico del autismo en Latinoamérica. Salud Cienc Med. 2021;1(2):14-25. Full text link

4. Catelli CLRQ, D'Antino MEF, Blascovi-Assis SM. Aspectos motores em indivíduos com transtorno do espectro autista: revisão de literatura. Cad Pos-Grad Disturb Desenvolv. 2016; 16(1):55-65. Full text link

5. Van Hecke R, Danneels M, Dhooge I, Van Waelvelde H, Wiersema JR, Deconinck FJA, et al. Vestibular function in children with neurodevelopmental disorders: a systematic review. J Autism Dev Disord. 2019;49(8):3328-50. DOI

6. Lloyd M, MacDonald M, Lord C. Motor skills of toddlers with autism spectrum disorders. Autism. 2013;17(2):133-46. DOI

7. Azevedo A, Gusmão M. A importância da fisioterapia motora no acompanhamento de crianças autistas. Rev Eletron Atualiza Saude. 2016;2(2):76-83. Full text link

8. Wilson RB, Vangala S, Elashoff D, Safari T, Smith BA. Using wearable sensor technology to measure motion complexity in infants at high familial risk for autism spectrum disorder. Sensors. 2021;21(2):616. DOI

9. Heidrich TE, Bastianel L, Gelain GM, Candotti CT. Content validity of an instrument for motor assessment of youth with autism. Fisioter Mov. 2022;35:e35135. DOI

10. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy. terminology. and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45. DOI

11. Sim J, Wright CC. The kappa statistic in reliability studies: use. interpretation. and sample size requirements. Phys Ther. 2005;85(3):257-68. DOI

12. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. Int J Nurs Stud. 2011;48(6):661-71. DOI

13. Gaya A, Garlipp DC, Silva MF, Moreira RB. Ciências do movimento humano: introdução à metodologia da pesquisa. Porto Alegre: Artmed; 2008. p. 241-303.

14. Bujang MA, Omar ED, Baharum NA. A review on sample size determination for cronbach's alpha test: a simple guide for researchers. Malays J Med Sci. 2018;25(6):85-99. DOI

15. Denegar CR, Ball DW. Assessing reliability and precision of measurement: An introduction to intraclass and standard error of measurement. J Sport Rehabil. 1993;2(1):35-42. DOI

16. Pivotto LR, Navarro IJRL, Candotti CT. Radiography and photogrammetry-based methods of assessing cervical spine posture in the sagittal plane: A systematic review with meta-analysis. Gait Posture. 2021;84:357-67. DOI

17. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. Hoboken: John Wiley & Sons; 2003.

18. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101-10. DOI

19. Allen KA, Bredero B, Van Damme T, Ulrich DA, Simons J. Test of gross motor development-3 (TGMD-3) with the use of visual supports for children with autism spectrum disorder: validity and reliability. J Autism Dev Disord. 2017;47(3):813-33. DOI

20. Colombo-Dougovito AM, Block ME, Zhang X, Strehli I. A multiple-method review of accommodations to gross motor assessments commonly used with children and adolescents on the autism spectrum. Autism. 2020;24(3):693-706. DOI

21. Gontijo KNS, Candotti CT, Feijó GS, Ribeiro LP, Loss LF. Dynamic evaluation of lower limbs joint alignment (MADAAMI) for dancers during the plié. Rev Bras Cienc Esporte. 2017; 39(2):148-59. DOI

22. Staples KL, MacDonald M, Zimmer C. Assessment of motor behavior among children and adolescents with autism spectrum disorder. Int Rev Res Dev Disabil. 2012;42(1):179-214. DOI

23. Leite IP, Diniz EFFS. Principais instrumentos para avaliar o desempenho motor em crianças com o transtorno de espectro autista: um estudo de revisão sistemática. Rev Assoc Bras Ativ Mot Adapt. 2022;23(1):35-52. DOI

24. Gallahue DL, Ozmun JD, Goodway JD. Compreendendo o desenvolvimento motor: bebês, crianças, adolescentes e adultos. 7 ed. Porto Alegre: Artmed; 2013. p. 273-90.

25. Rival C, Ceyte H, Olivier I. Developmental changes of static standing balance in children. Neurosci Lett. 2005;376(2):133-6. DOI

26. Stins JF, Emck C. Balance performance in autism: a brief overview. Front Phychol. 2018;9:901. DOI