children

MDPI

*Article*

# Reliability and Validity of the Mastication Observation and Evaluation Tool for Chewing Assessment in Children and Adolescents with Autism

**Ayşe Kübra Söyler [1,*], Ömer Faruk Yaşaroğlu [2], Nazan Öztürk [3] and Selen Serel Arslan [4]**

[1] Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Aydın Adnan Menderes University, Aydın 09010, Türkiye
[2] Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Harran University, Şanlıurfa 63300, Türkiye
[3] Söke Vocational School of Health Services, Aydın Adnan Menderes University, Aydın 09100, Türkiye
[4] Faculty of Physical Therapy and Rehabilitation, Hacettepe University, Ankara 06100, Türkiye
* Correspondence: ayse.kubra.sahan@adu.edu.tr

## Highlights

### What are the main findings?

- The Mastication Observation and Evaluation (MOE) tool demonstrated excellent intra-rater reliability (ICC = 0.96) and good inter-rater reliability (ICC = 0.79) for assessing chewing function in autism.
- MOE total scores showed a strong negative association with Karaduman Chewing Performance Scale (KCPS) scores ($r = -0.766$, $p < 0.001$), supporting the convergent validity of the MOE in relation to an established clinical measure of chewing performance.

### What are the implications of the main findings?

- The MOE may help clinicians characterize specific components of chewing dysfunction in autism beyond an overall classification of chewing performance, thereby contributing to clinical assessment.

## Abstract

Objective: The aim of this study was to evaluate the reliability and convergent validity of the Mastication Observation and Evaluation (MOE) tool in assessing chewing function in children and adolescents with autism (CAwA). Method: All participants were video-recorded while chewing a standardized biscuit. The videos were independently scored by two experienced physical therapists using the MOE tool. One of the physical therapists also rescored the videos after an interval of one week. Intra-rater reliability and inter-rater reliability were analyzed using the intraclass correlation coefficient (ICC). The Karaduman Chewing Performance Scale (KCPS) was used to evaluate the chewing performance levels of participants. The convergent validity of the MOE was examined by assessing its association with the KCPS. Results: This study was completed with 41 participants, with a mean age of 11.42 ± 3.82 years, of whom 85.4% were male. Average-measures intra-rater reliability ICC values for the MOE sub-items ranged from 0.75 to 0.97, and the total=score ICC was found to be 0.96. Average-measures inter-rater reliability ICC values ranged from 0.60 to 0.89 for the sub-items, and the total score ICC was found to be 0.79. A strong and negative correlation was found between the MOE total score and the KCPS score ($r = -0.766$; $p < 0.001$), supporting the convergent validity of MOE. Conclusions: This study demonstrates that the MOE has excellent intra-rater reliability and good overall

Check for updates

Academic Editor: Michael B. Brimacombe

Received: 10 September 2026
Revised: 25 October 2026
Accepted: 26 October 2026
Published: 27 October 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

inter-rater reliability for averaged ratings. The association with the KCPS provides evidence supporting its convergent validity for assessing chewing function in CAwA. Due to its multidimensional observational approach, the MOE may contribute to clinical assessment by providing information not only on overall chewing performance but also on affected components of chewing function.

**Keywords:** autism; mastication; chewing; psychometric validation; pediatric; feeding

## 1. Introduction

Autism is a neurodevelopmental condition involving persistent difficulties in social communication and social interaction, together with restricted or repetitive patterns of behavior, interests, or activities [1]. The increasing global prevalence of autism has presented significant challenges not only for affected individuals but also for their families [2]. Among the associated challenges, feeding-related problems, although common in children and adolescents with autism (CAwA), are often overlooked. These problems typically emerge in the early stages of development, often even before concerns related to autism are formally identified [3]. Previous studies have shown that children with autism experience more frequent and different feeding-related problems compared to their neurotypically developing peers and children with other developmental disorders [4–9]. The most common reported feeding-related problems in CAwA include picky eating, food refusal, chewing and swallowing difficulties, repetitive behaviors, and problematic feeding behaviors during mealtimes, such as closing the mouth, spitting, and keeping food in the mouth [5,8,9].

In particular, a previous study involving children with autism aged 4–12 years reported varying degrees of chewing difficulties in more than half of participants, as assessed using the Karaduman Chewing Performance Scale (KCPS) [8]. Chewing is a fundamental oral motor skill necessary for the safe and effective processing of solid foods and involves a complex interplay of jaw and tongue movements [10]. Chewing function develops progressively with the transition to complementary and solid foods and continues to mature as tongue–jaw coordination develops [11]. However, sensory, motor, and behavioral difficulties commonly observed in autism may negatively affect the development of chewing function and contribute to chewing difficulties [8,12]. Although feeding-related problems such as food avoidance and gagging may also have sensory or behavioral causes, chewing difficulties may adversely affect the processing of solid foods, potentially leading to the avoidance of food textures requiring chewing or to episodes of gagging or choking during attempts to swallow inadequately chewed food [12–15]. In autism, difficulties with solid foods may be interpreted as food refusal or selectivity, potentially resulting in underlying chewing impairments being overlooked [8,12]. Therefore, in CAwA, accurate and detailed assessment of chewing function is important not only for nutritional adequacy but also for swallowing safety and support of long-term oral motor development. Failing to identify the components of chewing disorders can make it difficult to plan appropriate food texture adjustments and targeted oral motor interventions to meet the child's needs. These problems can lead to malnutrition and developmental delays. They can also necessitate special meal preparation, thus placing a greater burden on caregivers, increasing stress levels, limiting social participation, and ultimately reducing the quality of life for both children and their families [13,16]. Furthermore, if left unaddressed, these feeding-related problems may persist into adulthood [17]. Therefore, early identification of chewing disorders through appropriate assessment methods is critical for adapting effective management strategies [18]. Understanding the nature of the chewing disorder may help in individualized dietary planning and the adaptation of food textures [15].

For chewing assessment, there are several objective and subjective methods. While objective methods including electromyography, motion analysis, particle size measurements, and two-color gum tests provide objective data on the physiological or mechanical aspects of chewing, they are often not applicable in CAwA due to the need for a high level of cooperation and standardized application [10]. Similarly, parent-reported questionnaires may indicate the presence of chewing disorders, but they cannot reveal which components of chewing function are affected. This highlights the need for reliable tools for assessing chewing function in autism in a structured and observational manner. However, most current studies on chewing assessment in autism rely on parent-reported questionnaires or screening tools, and often limit the assessment to a binary "present/absent" approach [19,20]. Current tools for assessing chewing function are limited in children; however, available measurement tools include the KCPS, the Test of Masticating and Swallowing Solids (TOMASS), and the MOE [21]. While the KCPS is a practical and rapid tool for evaluating chewing performance level, it does not provide detailed information on which components of chewing are affected [22]. On the other hand, TOMASS provides more objective data through quantitative parameters such as duration, number of chews, and number of swallows; however, it is affected by the type of food used and the participant's cooperation. Sensory deficits and behavioral problems, particularly in CAwA, can limit the applicability of performance-based tests like TOMASS. However, the MOE does not limit the chewing assessment to overall performance-based or time-based measurements; it addresses components such as lip, tongue and jaw movements, chewing time, food control, and fluency/coordination within a structured observation framework [23,24]. In this respect, it may complement existing assessment tools by providing component-level information beyond an overall classification of chewing performance.

Although the reliability and validity of the MOE have been previously evaluated in children with typical development and those with cerebral palsy, its reliability and validity have not yet been specifically investigated in CAwA [23–25]. Sensory and behavioral difficulties, variability in cooperation and oral motor development, and differences in the visibility of chewing movements during assessment may influence MOE item scoring and inter-rater agreement in this population. Therefore, measurement properties established in other pediatric populations may not directly apply to CAwA. Despite this gap, a recent study utilized the MOE to assess chewing performance in children with autism without evaluating reliability or validity [26]. This indicates that the MOE has begun to be used in CAwA despite the lack of established reliability and validity evidence in this population. Therefore, the aim of this study was to evaluate the inter-rater and intra-rater reliability and convergent validity of the MOE in CAwA.

## 2. Method

Ethical approval for the study was granted by the Aydın Adnan Menderes University Non-Interventional Clinical Research Ethics Committee (Approval Number: E-15189967-050.04-603792). The study procedures complied with the principles outlined in the Declaration of Helsinki. Data were collected between September 2024 and May 2025 from children diagnosed with autism who presented to the Department of Physiotherapy and Rehabilitation at Aydın Adnan Menderes University and were recruited using a consecutive sampling method.

Informed written consent was obtained from the parents of all participants after they were fully informed about the purpose and procedures of the study. The participants' willingness to participate was respected throughout the assessment. Inclusion criteria were as follows: a formal diagnosis of autism established by child and adolescent psychiatrists according to DSM-5 criteria, and age between 5 and 18 years. Participants with additional

neurological or neurodevelopmental disorders (e.g., cerebral palsy, epilepsy, or diagnosed intellectual disability), a history of head or neck surgery affecting oral motor function, or dependence on enteral feeding, as well as participants who declined to chew the standardized biscuit during the assessment or who accepted the biscuit but were unable to complete the standardized three-evaluable-bite assessment, were excluded from the study [8,23,26]. These medical exclusion criteria were applied to reduce clinical heterogeneity when evaluating the reliability and convergent validity of the MOE, specifically in CAwA. Therefore, the study sample was limited to participants who were able and willing to complete the standardized chewing task. The autism diagnosis and other medical information relevant to eligibility were verified through medical record review. The participant flow diagram is presented in Figure 1.

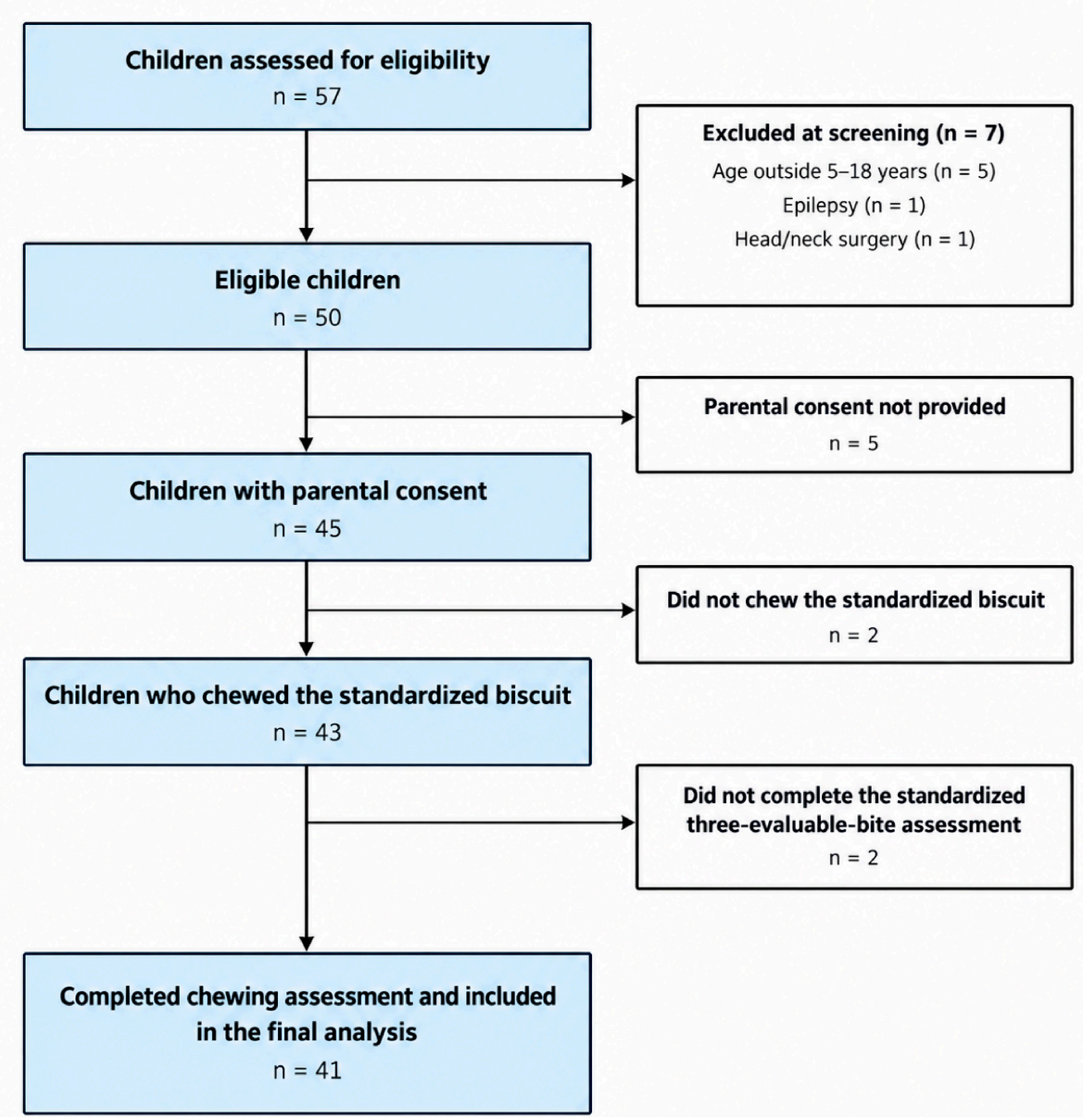


**Figure 1.** Flow diagram.

The sample size was planned pragmatically with reference to methodological recommendations for psychometric validation studies, in which sample size is commonly determined using the participant-to-item ratio [27]. Previous methodological literature suggests that participant-to-item ratios ranging from 5:1 to 10:1 are commonly used in

psychometric validation studies, although there is no absolute rule regarding the required sample size and recommendations may vary depending on the characteristics of the instrument and the study design [27,28]. Since the MOE consists of eight items, the recruitment target was 40 participants. Ultimately, 41 CAwA were included in the study, with the participant-to-item ratio used as a pragmatic recruitment guide rather than a formal sample-size calculation.

### *2.1. Outcome Measures and Data Collection Procedures*

Participants' age, height, and weight were recorded as demographic data. Feeding-related characteristics such as the age of transition to additional food and solid food and the age of first tooth eruption were obtained via parental report and noted in months, as described in previous studies [8,24]. Observational evaluation of oral structural problems, open-mouth posture and high-arched palate, was performed and recorded as 'present' or 'absent' [8].

To evaluate chewing function, participants were seated in a chair and instructed to chew a standardized biscuit while being video-recorded. Caregivers remained beside the children to provide reassurance and, when necessary, repeated this instruction. Their involvement was limited to this supportive role and did not constitute a separate assessment procedure. A commercially available biscuit (Ülker Pötibör Biscuit, Ülker, Istanbul, Türkiye; approximately 7.0 × 5.2 cm and 5.8 g per biscuit) was used as the standardized test food. The biscuit contained sugar, wheat flour, vegetable oil, milk, whey powder, invert sugar syrup, salt, and leavening agents. The videos were recorded from the front and from a fixed angle, allowing the participants' facial and mouth movements to be clearly seen [8,23]. The assessment was based on three evaluable bites, consistent with the three-bite observation approach described by Remijn et al. [23]. Instead of using a fixed recording duration, video recording continued until this criterion was met. If a participant could not complete the assessment, recording was stopped, and the participant was not included in the final analysis. Each chewing video was subsequently scored using the KCPS and the MOE tool. The KCPS is a valid and reliable measurement tool for classifying chewing performance in the pediatric population. The scale classifies chewing function into five different levels from 0 to 4, with higher levels indicating greater functional impairment. 'Level 0' describes normal chewing function, in which the child can retain food in the mouth, transfer it to the molar area, break it down to form a bolus, and swallow safely. 'Level 1' describes difficulties in bolus formation including prolonged chewing time, insufficient chewing force, and food/saliva loss. 'Level 2' indicates that the child can begin chewing but cannot keep food in the molar area, and bolus formation is impaired. 'Level 3' means that the child can bite but cannot chew, and 'level 4' indicates that the child is not able to bite and chew [22]. On the other hand, the MOE tool is an observational scale developed to systematically assess the components of chewing. The Turkish version of the MOE was used in the present study. It consists of eight items including tongue protrusion, lateral tongue movement, squashing or sucking tongue movement, jaw movement, chewing duration, food loss or saliva, number of swallows, and fluency/coordination. Each item is scored on a 4-point ordinal scale from 1 (very inappropriate) to 4 (very appropriate). The total score ranges from 8 to 32, with higher scores indicating better chewing performance and fewer impairments in chewing function [23–25]. Item scores were assigned according to established item-specific scoring criteria [23,24]. All eight MOE items were scored for every participant included in the final analysis; there were no missing item-level scores.

### 2.2. Validity and Reliability

The intra-rater reliability and inter-rater reliability were investigated as reliability measures. Two physical therapists (Rater 1 and Rater 2), each with at least seven years of clinical experience in swallowing disorders, independently scored all chewing video recordings using the MOE. The study videos were evaluated separately, and the raters were blinded to each other's scores. The scoring process was completed independently, without any discussion of the study videos between the raters. To examine the intra-rater reliability of the MOE, Rater 2 rescored the same video recordings after a one-week interval without access to the initial scores.

A gold-standard reference instrument is generally needed to evaluate the criterion validity of a measurement instrument. However, there is no accepted gold standard for evaluating chewing function in the pediatric population. The KCPS was selected because it is widely used in clinical practice and has been validated for evaluating chewing performance in children [18,24]. Therefore, convergent validity was examined by investigating the association between the MOE and KCPS scores. For this analysis, Rater 1 first scored and recorded the KCPS results and subsequently scored the same video recordings using the MOE after a one-week interval without access to the previously recorded KCPS scores, to minimize potential recall bias. For the convergent validity analysis, the KCPS scores were correlated with Rater 1's MOE total score.

### 2.3. Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics version 25 (IBM Corp., Armonk, NY, USA). Variable distributions were assessed through histograms, probability plots, and the Shapiro–Wilk test. Inter-rater and intra-rater reliability of the MOE was examined using average-measures intraclass correlation coefficients (ICCs) derived from a two-way random-effects model with absolute agreement. The model was selected to account for variability across participants and raters or scoring occasions, while absolute agreement was used because differences in actual MOE scores were relevant. ICC estimates were classified as poor (<0.50), moderate (0.50–0.75), good (0.75–0.90), or excellent (>0.90) [29]. As the MOE total scores did not demonstrate a normal distribution, Spearman's rank-order correlation coefficient was used to examine the association between KCPS and MOE total scores. The 95% confidence interval for Spearman's correlation coefficient was estimated using 20,000 paired bootstrap resamples and the percentile method. Correlation strength was classified as negligible (0.00–0.10), weak (0.10–0.39), moderate (0.40–0.69), strong (0.70–0.89), or very strong (0.90–1.00) [30]. A significance level of $p < 0.05$ was considered statistically significant.

To account for the four-category ordinal structure of the MOE items, item-level agreement was additionally assessed using linearly weighted Cohen's kappa, exact agreement proportions, and category-specific agreement percentages. Inter-rater analyses compared the two raters' initial scores, whereas intra-rater analyses compared the same rater's initial and one-week repeat scores of the same videos. Linear agreement weights were defined as $w_{ij} = 1 - |i - j|/3$. The 95% confidence intervals for weighted kappa were estimated using 20,000 paired bootstrap resamples and the percentile method. Resamples yielding undefined kappa values were excluded from percentile estimation. Category-specific agreement was calculated as $2n/(n_{+} + n_{+}) \times 100$, with categories unused in both ratings considered not estimable [31].

Absolute disagreement in MOE total scores was assessed separately for inter-rater and intra-rater comparisons. Signed differences were calculated as Rater 1 minus Rater 2 for inter-rater agreement and initial minus repeated ratings for intra-rater agreement. Mean differences and their 95% confidence intervals were calculated using the paired-

difference standard error and the t distribution. Mean absolute differences, exact total-score agreement proportions, and observed difference ranges were also calculated. Difference-versus-mean plots and the distributions of paired differences were examined to assess the appropriateness of conventional Bland–Altman limits of agreement. Fixed normal-theory limits were not adopted as the primary summary because the difference distributions departed from normality and, particularly for intra-rater comparisons, their variability depended on score magnitude [32].

## 3. Results

This study included forty-one CAwA, 85.4% of whom were male. Table 1 presents the descriptive and feeding-related characteristics of the participants. The mean age of the participants was 11.42 ± 3.82 years. The age distribution was as follows: 6–9 years (n = 12, 29.3%), 10–14 years (n = 20, 48.8%), and 15–18 years (n = 9, 22.0%). One participant was 18 years old. Regarding oral characteristics, 63.4% (n = 26) of the children presented with a high-arched palate, and 9.8% (n = 4) presented with an open mouth. Regarding chewing assessment, the KCPS distribution is presented in Table 2 and the MOE results are presented in Table 3. The maximum possible MOE total score of 32 was observed in 5 (12.2%), 8 (19.5%), and 10 (24.4%) children in the Rater 1, Rater 2, and Rater 2 repeat assessments, respectively, whereas none of the children obtained the minimum possible score of 8.

**Table 1.** Descriptive and feeding-related characteristics of children and adolescents with autism.

| **Demographic Characteristics** | **Mean ± SD (Min–Max)** |
|---|---|
| Age (years) | 11.42 ± 3.82 (6–18) |
| Height (cm) | 154.65 ± 20.07 (110–183) |
| Weight (kg) | 55.53 ± 25.62 (16–89) |
| Gender | Number (%) |
| Boy | 35 (85.4) |
| Girl | 6 (14.6) |
| **Parameters Related to Oral Motor Development** | **Mean ± SD (Min–Max)** |
| Transition time to additional food (month) | 6.76 ± 1.92 (4–13) |
| Transition time to solid food (month) | 10.08 ± 4.13 (6–24) |
| Initial teething time (month) | 6.19 ± 0.90 (4–8) |
| **Oral Structural Parameters** | **Number (%)** |
| High-arched palate | 26 (63.4) |
| Open mouth | 4 (9.8) |

**Table 2.** Chewing evaluation results according to KCPS classification.

| **Karaduman Chewing Performance Scale (KCPS)** | **Number (%)** |
|---|---|
| Level 0: Normal chewing function | 9 (22.0) |
| Level 1: The child chews, but there are some difficulties in food transition to bolus | 27 (65.9) |
| Level 2: The child starts to chew, but he/she cannot hold the food in the molar area | 4 (9.8) |
| Level 3: The child bites but cannot chew | 1 (2.4) |
| Level 4: The child cannot bite and chew | 0 |

**Table 3.** Item-level and total MOE scores for the three separate ratings.

| MOE Item | Rater 1<br>Mean ± SD (Min–Max) | Rater 2<br>Mean ± SD (Min–Max) | Rater 2 Repeat<br>Mean ± SD (Min–Max) |
|---|---|---|---|
| Item 1: Tongue protrusion | 3.51 ± 0.93 (1–4) | 3.39 ± 1.14 (1–4) | 3.34 ± 1.15 (1–4) |
| Item 2: Lateral tongue movement | 3.17 ± 0.67 (2–4) | 3.27 ± 0.74 (1–4) | 3.29 ± 0.81 (1–4) |
| Item 3: Squashing or sucking movement | 2.95 ± 0.80 (1–4) | 3.05 ± 1.05 (1–4) | 3.10 ± 1.09 (1–4) |
| Item 4: Jaw movement | 3.20 ± 0.78 (1–4) | 3.44 ± 0.67 (1–4) | 3.44 ± 0.67 (1–4) |
| Item 5: Chewing duration | 3.63 ± 0.66 (2–4) | 3.49 ± 0.81 (2–4) | 3.46 ± 0.84 (1–4) |
| Item 6: Loss of food or saliva | 3.85 ± 0.36 (3–4) | 3.85 ± 0.36 (3–4) | 3.78 ± 0.42 (3–4) |
| Item 7: Number of swallows | 3.80 ± 0.40 (3–4) | 3.88 ± 0.40 (2–4) | 3.78 ± 0.52 (2–4) |
| Item 8: Fluency/coordination | 2.54 ± 0.95 (1–4) | 2.93 ± 0.96 (1–4) | 3.02 ± 0.94 (1–4) |
| Total score | 26.66 ± 3.31 (19–32) | 27.29 ± 3.71 (20–32) | 27.22 ± 4.14 (18–32) |

Intra-rater and inter-rater reliability results are presented in Table 4. The average-measures ICC for the MOE total score was 0.96 (95% CI: 0.93–0.98), indicating excellent repeatability of scoring the same video recordings. MOE item-level ICCs ranged from 0.75 to 0.97, representing good to excellent reliability. The highest intra-rater agreement was found for 'lateral tongue movement' (Item 2; ICC = 0.97) and 'fluency/coordination' (Item 8; ICC = 0.96). The average-measures ICC for the total score was 0.79 (95% CI: 0.60–0.89), indicating good inter-rater agreement for ratings averaged across the two raters. However, its confidence interval was relatively wide, and the confidence intervals for the three lowest-ICC MOE items extended into the poor-reliability range, with lower limits ranging from 0.28 to 0.31. MOE item-level ICCs ranged from 0.60 to 0.89 for inter-rater reliability and showed the highest inter-rater reliability for 'loss of food or saliva' (Item 6; ICC = 0.89). The lowest inter-rater reliability values were observed for jaw movement (Item 4; ICC = 0.60), lateral tongue movement (Item 2; ICC = 0.63), and squashing or sucking movement (Item 3; ICC = 0.63). Spearman correlation analysis revealed a strong negative association between the MOE total score and the KCPS ($r = -0.766$; 95% bootstrap CI: $-0.872$ to $-0.594$; $p < 0.001$). The distribution of MOE total scores across KCPS categories is presented in Figure 2.

**Table 4.** Intra-rater and inter-rater reliability of MOE.

| Mastication Observation and Evaluation (MOE) | Intra-Rater | | Inter-Rater | |
|---|---|---|---|---|
| | ICC | 95% CI | ICC | 95% CI |
| Item 1: Tongue protrusion | 0.93 | 0.87–0.96 | 0.80 | 0.63–0.89 |
| Item 2: Lateral tongue movement | 0.97 | 0.94–0.98 | 0.63 | 0.31–0.80 |
| Item 3: Squashing or sucking movement | 0.91 | 0.82–0.95 | 0.63 | 0.31–0.81 |
| Item 4: Jaw movement | 0.94 | 0.89–0.97 | 0.60 | 0.28–0.79 |
| Item 5: Chewing duration | 0.82 | 0.66–0.90 | 0.72 | 0.47–0.85 |
| Item 6: Loss of food or saliva | 0.75 | 0.54–0.87 | 0.89 | 0.80–0.94 |
| Item 7: Number of swallows | 0.80 | 0.63–0.89 | 0.77 | 0.57–0.88 |
| Item 8: Fluency/coordination | 0.96 | 0.92–0.98 | 0.73 | 0.55–0.87 |
| Total score | 0.96 | 0.93–0.98 | 0.79 | 0.60–0.89 |

ICC: intraclass correlation coefficient; CI: confidence interval. All reported ICC values are average-measures estimates based on two ratings.

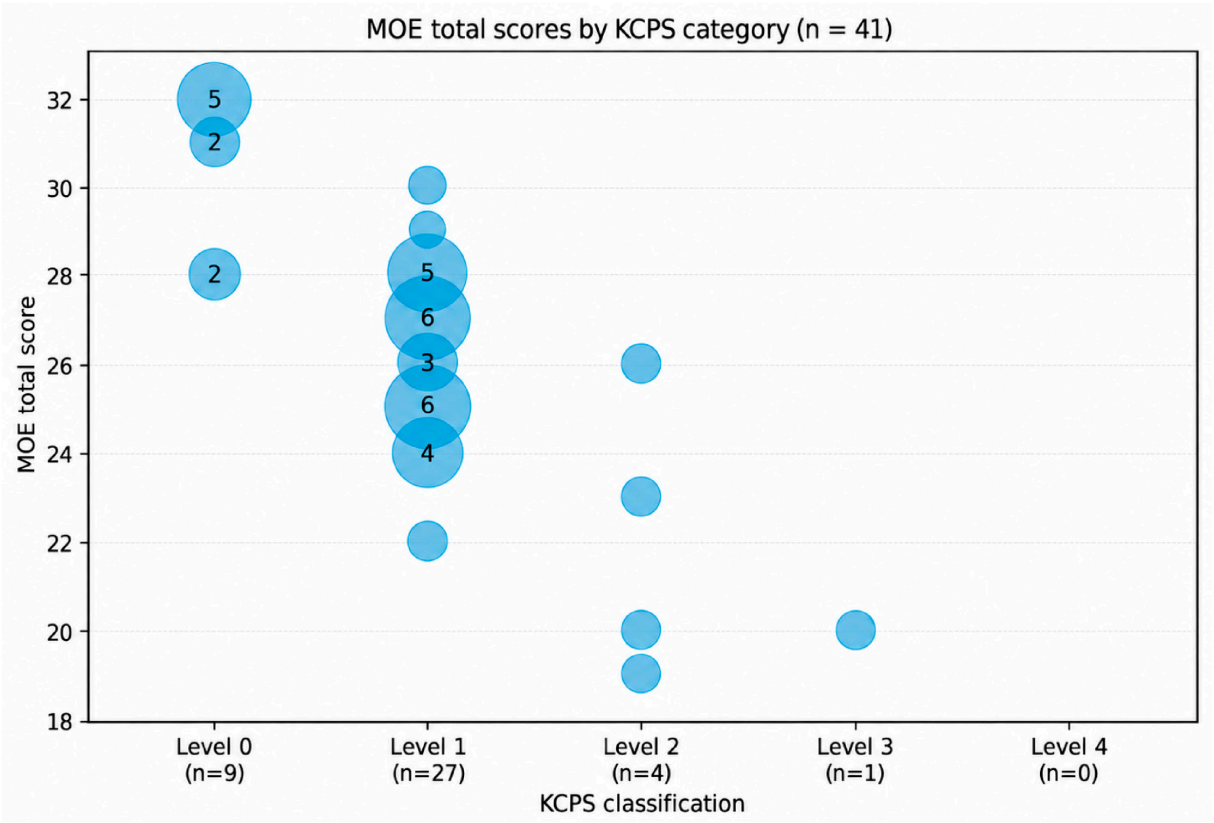


**Figure 2.** Distribution of MOE total scores across KCPS categories. Each bubble represents one or more children with identical MOE total scores and KCPS classifications. Bubble size is proportional to the number of overlapping observations, with numbers inside bubbles indicating their frequencies. The association between the MOE and KCPS was assessed using Spearman's rank correlation coefficient ($r = -0.766$; 95% bootstrap CI: $-0.872$ to $-0.594$; $p < 0.001$).

Supplementary analyses of the four-category MOE items yielded linearly weighted kappa values of 0.344–0.805 for inter-rater and 0.586–0.901 for intra-rater agreement. Exact agreement ranged from 43.9% to 95.1% and from 73.2% to 92.7%, respectively. The lowest inter-rater weighted kappa was observed for squashing or sucking movement, whereas the highest was observed for loss of food or saliva. For intra-rater agreement, the lowest weighted kappa was observed for chewing duration and the highest for lateral tongue movement. Category-specific agreement varied across items and response categories. Item- and category-specific results are reported in Supplementary Table S1. Across the 41 paired observations, the mean inter-rater total-score difference (Rater 1 − Rater 2) was −0.63 points (95% CI: −1.56 to 0.29), and the mean intra-rater difference (Rater 2 − Rater 2 repeat) was 0.07 points (95% CI: −0.41 to 0.56). Neither confidence interval excluded zero, providing no clear evidence of a systematic mean score difference. Mean absolute differences were 2.00 and 0.90 points, respectively. Scores agreed exactly in 11/41 inter-rater assessments (26.8%) and 21/41 intra-rater assessments (51.2%); observed signed differences ranged from −10 to 7 and −5 to 4 points, respectively. Supplementary Figure S1 shows the corresponding difference-versus-mean plots.

## 4. Discussion

This study evaluated the reliability and convergent validity of the MOE to assess chewing function in CAwA. In terms of reliability analyses, the MOE demonstrated high intra-rater agreement and good overall inter-rater reliability. In convergent validity analysis,

the findings showed a strong association between the MOE and KCPS, which supports the convergent validity of the MOE in CAwA.

In the present study, intra-rater reliability reflected the repeatability of scoring the same video recordings by the same evaluator at different times. The average-measures ICC value of 0.96 obtained for the MOE total score falls within the "excellent" reliability limits defined in the literature [29]. The high intra-rater reliability achieved even with items involving more complex oral motor components such as lateral tongue movement, squashing or sucking movement, and jaw movement suggests that averaged ratings of these parameters are repeatable when the same evaluator reviews the same recordings. While studies on the validity and reliability of the MOE are limited, the current findings seem to be consistent with the literature. Indeed, a validation study including children with typical development and with cerebral palsy reported high intra-rater reliability (ICC: 0.89–0.97) and inter-rater reliability (ICC: 0.86–0.94) [24]. These findings support the repeatability of MOE scoring by the same evaluator.

Inter-rater reliability is a fundamental psychometric property that indicates the extent to which a measurement tool yields consistent and objective results when used by different raters. In clinical settings, it is expected that scores obtained from an assessment tool produce consistent results thanks to structured and standardized scoring criteria [33]. Especially in observational measurements such as the assessment of chewing function, differences in interpretation between raters can lead to measurement error. Therefore, a high level of inter-rater agreement is an important indicator supporting the objectivity and clinical applicability of the measurement tool. In our study, the MOE total-score ICC of 0.79 indicated good reliability for scores averaged across the two raters [29]. These findings suggest that averaged MOE scores may be reliable in the assessment of chewing function in CAwA, while item-level differences in agreement should be considered when interpreting scores. When inter-rater reliability is examined at the item level, some differences in agreement are observed among the different items of the MOE. The inter-rater agreement was higher for some items such as tongue protrusion, chewing duration, loss of food or saliva, number of swallows, and fluency/coordination. In contrast, it is noteworthy that inter-rater agreement was relatively lower in lateral tongue movement, squashing or sucking movement, and jaw movement items. These are consistent with the results reported in the original development and validation studies of the MOE [23,25]. Studies by Remijn et al. reported that inter-rater reliability varied among items in both typically developing children and those with cerebral palsy, and was particularly lower for items involving intraoral processes such as lateral tongue movement, squashing or sucking movement, and jaw movement. Furthermore, these studies showed that inter-rater reliability could vary not only depending on the item being assessed but also depending on the type of solid food used (bread and biscuits) and the rater's level of experience. The authors noted that a significant portion of these differences were due to the interpretation of the appropriateness of the behavior (e.g., the distinction between scores 3 and 4) and the varying visibility of intraoral movements [23,25]. Although jaw movement is an externally observable component of chewing, this item in the MOE is scored not only according to the presence of mandibular movement but also according to the adequacy and frequency of mandibular movement in the expected vertical and slight horizontal directions during chewing. Therefore, distinguishing between adjacent scoring categories (e.g., scores 3 and 4, reflecting occasional versus regular or constant movement from the midline) may be challenging for raters. This may partly explain the relatively lower inter-rater agreement observed for the jaw movement item. However, evidence regarding the performance of the MOE across different clinical populations remains limited. Future studies involving

additional pediatric populations may help to better clarify the potential reasons for the lower inter-rater agreement observed for the jaw movement item.

Although the inter-rater reliability observed in the present study appeared to be slightly lower than that reported in previous MOE validation studies conducted in children with typical development and with cerebral palsy, direct comparisons should be interpreted with caution because the available studies were performed in different clinical populations [23–25]. CAwA often exhibit behavioral variability and inconsistent cooperation during chewing assessments, which may complicate the observation and interpretation of subtle oral motor behaviors. Greater heterogeneity in oral motor performance among CAwA may also contribute to differences in observational scoring, even when standardized scoring criteria are used. These factors may contribute to the lower inter-rater agreement observed for some MOE items. The concentration of some item scores in higher response categories may also influence agreement estimates. Although the MOE provides standardized scoring criteria, raters may still differ in their interpretation of subtle oral motor behaviors, particularly when distinguishing between adjacent scoring categories. Similar observations have been reported in a previous MOE study, in which less experienced raters demonstrated slightly lower agreement than expert raters despite standardized training, particularly for more complex chewing components [25]. Taken together, these findings suggest high intra-rater agreement when the same video recordings are rescored, although agreement between raters varies across individual items. Therefore, while the overall good inter-rater reliability of averaged MOE total-score ratings supports its potential use as a standardized observational tool in clinical assessment, the lower agreement and uncertainty observed for some items, as well as occasional substantial differences in total scores between raters, should be considered when interpreting individual-component and total scores.

In validity analyses, a strong and significant negative correlation was found between the total MOE score and the KCPS, both of which were scored by the same assessor using the same video recordings. This indicates that the two instruments reflect chewing performance in an inverse but consistent manner, providing preliminary evidence of convergent validity for the MOE. This negative association was expected because higher MOE scores indicate better chewing performance, whereas higher KCPS scores indicate poorer chewing performance. Also, the strong association between the MOE and KCPS suggests that the two instruments assess related aspects of chewing function. This finding is similar to the very strong MOE–KCPS association reported in a study that examined the MOE in children with typical development and with cerebral palsy [24]. This similarity observed in the current study further supports the convergent validity evidence for the MOE in CAwA.

In addition, the mean total MOE score of 26.66 out of a maximum possible score of 32 was comparable to the mean score of 26.63 ± 4.23 reported by Shalini and Swapna (2025) in children with autism [26]. However, the mean MOE score alone cannot determine clinical severity in the absence of validated severity thresholds. According to the KCPS classification, most participants were classified as level 1 (65.9%). This distribution suggests that relatively mild chewing difficulties predominated in the present sample, consistent with previous findings in children with autism [8]. Nevertheless, the limited representation of participants with more severe chewing difficulties in the present sample limits conclusions about the reliability and convergent validity of the MOE in CAwA. Beyond the total score, the MOE also provides information on individual components of chewing function. This item-level assessment may provide more detailed information about chewing function than the total score alone. Another important practical feature of the MOE is its observational nature, which allows chewing function to be assessed without specialized instrumental equipment, although the procedure requires video recording and scoring of the recordings. However, chewing assessment in CAwA presents inherent challenges due to behavioral and sensory

problems. Some participants may not accept standard solid foods, may not take sufficient bites, or may not remain in the appropriate position during video recording; this can make it difficult to provide consistent observation conditions for each child. These factors may limit the applicability of the standardized assessment and should be considered when interpreting the findings, particularly as children unable to complete the standardized task were excluded from the analysis. Furthermore, the findings may not be generalizable to CAwA with the medical conditions specified in the exclusion criteria, and the single-center design may limit their generalizability. In addition, autism symptom severity was not assessed using a standardized instrument in the present study. Therefore, it remains unclear whether the observed reliability and convergent validity findings are generalizable across different levels of autism symptom severity. The broad age range of 6–18 years may also encompass developmental and behavioral differences relevant to chewing assessment, but age-stratified reliability analyses were not performed. In addition, the absence of a typically developing control group precluded the evaluation of discriminant validity. Lastly, although chewing function was evaluated using the KCPS in the present study, future studies may incorporate additional clinical information to further characterize this population. Future studies may consider examining the MOE across different food textures, assessing its reliability among clinicians with varying levels of experience, and exploring its responsiveness and the clinical significance of changes in MOE scores in independent samples.

In conclusion, the results of this study support the reliability of the MOE and provide preliminary evidence of convergent validity for assessing chewing function in CAwA. The MOE's structured and multidimensional observational approach may contribute to clinical assessment by providing information not only on overall chewing performance but also on the affected components of chewing function. This component-level information may help characterize different components of chewing difficulties among CAwA, even when their overall chewing performance levels are similar.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/children13111481/s1. Table S1: Ordinal item agreement for the Mastication Observation and Evaluation tool in children with autism. Figure S1. Inter-rater and intra-rater agreement in MOE total scores.

**Author Contributions:** Conceptualization, A.K.S. and S.S.A.; Methodology, A.K.S., Ö.F.Y., N.Ö. and S.S.A.; Formal Analysis, N.Ö.; Investigation, A.K.S. and Ö.F.Y.; Data Curation, N.Ö.; Writing—Original Draft, A.K.S. and Ö.F.Y.; Writing—Review and Editing, S.S.A.; Supervision, S.S.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Aydın Adnan Menderes University Faculty of Health Sciences Non-Interventional Clinical Research Ethics Committee (Approval Code: E-15189967-050.04-603792, Protocol No.: 2024/40, Decision No.: 01, Approval Date: 9 September 2024).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical considerations related to the protection of the confidentiality of pediatric participants.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. American Psychiatric Association (Ed.) *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*; American Psychiatric Association: Washington, DC, USA, 2022. [CrossRef]

2. Lord, C.; Elsabbagh, M.; Baird, G.; Veenstra-Vanderweele, J. Autism spectrum disorder. *Lancet* **2018**, *392*, 508–520. [CrossRef] [PubMed]
3. Emond, A.; Emmett, P.; Steer, C.; Golding, J. Feeding symptoms, dietary patterns, and growth in young children with autism spectrum disorders. *Pediatrics* **2010**, *126*, e337–e342. [CrossRef] [PubMed]
4. Provost, B.; Crowe, T.K.; Osbourn, P.L.; McClain, C.; Skipper, B.J. Mealtime behaviors of preschool children: Comparison of children with autism spectrum disorder and children with typical development. *Phys. Occup. Ther. Pediatr.* **2010**, *30*, 220–233. [CrossRef] [PubMed]
5. Mayes, S.D.; Zickgraf, H. Atypical eating behaviors in children and adolescents with autism, ADHD, other disorders, and typical development. *Res. Autism Spectr. Disord.* **2019**, *64*, 76–83. [CrossRef]
6. Malhi, P.; Venkatesh, L.; Bharti, B.; Singhi, P. Feeding problems and nutrient intake in children with and without autism: A comparative study. *Indian J. Pediatr.* **2017**, *84*, 283–288. [CrossRef] [PubMed]
7. Yeung, S.S.; Chan, R.; Li, L.; Chan, D.; Leung, J.; Leung, T.F. Eating behaviors and diet quality in Chinese preschoolers with and without autism spectrum disorder: A case–control study. *J. Pediatr.* **2021**, *237*, 258–266.e5. [CrossRef] [PubMed]
8. Şahan, A.K.; Öztürk, N.; Demir, N.; Karaduman, A.A.; Serel Arslan, S. A comparative analysis of chewing function and feeding behaviors in children with autism. *Dysphagia* **2021**, *36*, 993–998. [CrossRef] [PubMed]
9. Pinto-Silva, R.; Nunes Costa, A.M.; Tello-Rodrigues, I. Feeding problems in children with autism spectrum disorders: A systematic review. *Speech Lang. Hear.* **2023**, *26*, 130–141. [CrossRef]
10. Van der Bilt, A. Assessment of mastication with implications for oral rehabilitation: A review. *J. Oral Rehabil.* **2011**, *38*, 754–780. [CrossRef] [PubMed]
11. Le Révérend, B.J.; Edelson, L.R.; Loret, C. Anatomical, functional, physiological and behavioural aspects of the development of mastication in early childhood. *Br. J. Nutr.* **2014**, *111*, 403–414. [CrossRef] [PubMed]
12. Hyman, S.L.; Levy, S.E.; Myers, S.M.; Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics; Kuo, D.Z.; Apkon, S.; Davidson, L.F.; Ellerbeck, K.A.; Foster, J.E.A.; Noritz, G.H.; et al. Identification, Evaluation, and Management of Children with Autism Spectrum Disorder. *Pediatrics* **2020**, *145*, e20193447. [CrossRef] [PubMed]
13. Rogers, L.G.; Magill-Evans, J.; Rempel, G.R. Mothers' challenges in feeding their children with autism spectrum disorder—Managing more than just picky eating. *J. Dev. Phys. Disabil.* **2012**, *24*, 19–33. [CrossRef]
14. Cermak, S.A.; Curtin, C.; Bandini, L.G. Food selectivity and sensory sensitivity in children with autism spectrum disorders. *J. Am. Diet. Assoc.* **2010**, *110*, 238–246. [CrossRef] [PubMed]
15. Goday, P.S.; Huh, S.Y.; Silverman, A.; Lukens, C.T.; Dodrill, P.; Cohen, S.S.; Delaney, A.L.; Feuling, M.B.; Noel, R.J.; Gisel, E.; et al. Pediatric feeding disorder: Consensus definition and conceptual framework. *J. Pediatr. Gastroenterol. Nutr.* **2019**, *68*, 124–129. [CrossRef] [PubMed]
16. Thullen, M.; Bonsall, A. Co-parenting quality, parenting stress, and feeding challenges in families with a child diagnosed with autism spectrum disorder. *J. Autism Dev. Disord.* **2017**, *47*, 878–886. [CrossRef] [PubMed]
17. Bayoumi, S.C.; Halkett, A.; Miller, M.; Hinshaw, S.P. Food selectivity and eating difficulties in adults with autism and/or ADHD. *Autism* **2025**, *29*, 1497–1509. [CrossRef] [PubMed]
18. Serel Arslan, S.; Aydın, G.; Alemdaroğlu, İ.; Tunca Yılmaz, Ö.; Karaduman, A.A. Reliability and validity of the Karaduman Chewing Performance Scale in paediatric neuromuscular diseases: A system for classification of chewing disorders. *J. Oral Rehabil.* **2018**, *45*, 526–531. [CrossRef] [PubMed]
19. Gal, E.; Gal-Mishael, R.; Vissoker, R.E.; Hedley, D.; Bury, S.M.; Stolar, O. Eating challenges in children with autism spectrum disorder: Development and validation of the "Aut-Eat" Questionnaire (AEQ). *J. Autism Dev. Disord.* **2022**, *52*, 811–822. [CrossRef] [PubMed]
20. Lukens, C.T.; Linscheid, T.R. Development and validation of an inventory to assess mealtime behavior problems in children with autism. *J. Autism Dev. Disord.* **2008**, *38*, 342–352. [CrossRef] [PubMed]
21. Marshall, J.; Gosa, M.M.; Dodrill, P. A systematic review of clinical assessment tools that evaluate the feeding skills domain for paediatric feeding disorder. *Int. J. Speech-Lang. Pathol.* **2025**, *28*, 652–662. [CrossRef] [PubMed]
22. Serel Arslan, S.; Demir, N.; Barak Dolgun, A.; Karaduman, A.A. Development of a new instrument for determining the level of chewing function in children. *J. Oral Rehabil.* **2016**, *43*, 488–495. [CrossRef] [PubMed]
23. Remijn, L.; Speyer, R.; Groen, B.E.; van Limbeek, J.; Nijhuis-van der Sanden, M.W. Validity and reliability of the Mastication Observation and Evaluation (MOE) instrument. *Res. Dev. Disabil.* **2014**, *35*, 1551–1561. [CrossRef] [PubMed]
24. Serel Arslan, S.; Demir, N.; Karaduman, A.A. Turkish Version of the Mastication Observation and Evaluation (MOE) Instrument: A Reliability and Validity Study in Children. *Dysphagia* **2020**, *35*, 328–333. [CrossRef] [PubMed]
25. Remijn, L.; Speyer, R.; Groen, B.E.; Holtus, P.C.M.; Van Limbeek, J.; Nijhuis-Van der Sanden, M.W.G. Assessment of mastication in healthy children and children with cerebral palsy: A validity and consistency study. *J. Oral Rehabil.* **2013**, *40*, 336–347. [CrossRef] [PubMed]

26. Shalini, M.; Swapna, N. The Chewing Challenge: Linking Oral Motor and Oral Sensory Skills to Feeding Difficulties in Autism. *J. Indian Speech Lang. Hear. Assoc.* **2025**, *39*, 63–67. [CrossRef] [PubMed]
27. Stefana, A.; Damiani, S.; Granziol, U.; Provenzani, U.; Solmi, M.; Youngstrom, E.A.; Fusar-Poli, P. Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Front. Psychol.* **2025**, *15*, 1494261. [CrossRef] [PubMed]
28. Tsang, S.; Royse, C.F.; Terkawi, A.S. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J. Anaesth.* **2017**, *11*, S80–S89. [CrossRef] [PubMed]
29. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. Erratum in *J. Chiropr. Med.* **2017**, *16*, 346. [CrossRef] [PubMed]
30. Schober, P.; Boer, C.; Schwarte, L.A. Correlation coefficients: Appropriate use and interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. [CrossRef] [PubMed]
31. de Vet, H.C.; Mullender, M.G.; Eekhout, I. Specific agreement on ordinal and multiple nominal outcomes can be calculated for more than two raters. *J. Clin. Epidemiol.* **2018**, *96*, 47–53. [CrossRef] [PubMed]
32. Bland, J.M.; Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [CrossRef]
33. Bobak, C.A.; Barr, P.J.; O'Malley, A.J. Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. *BMC Med. Res. Methodol.* **2018**, *18*, 93. [CrossRef] [PubMed]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.