

Research Paper: Test-Retest and Inter-Rater Reliability Study of the Schedule for Oral-Motor Assessment in Persian Children



Manijeh Zarei Mahmood Abadi¹, Fariba Yadegari^{1*}, Mehrzad Mehdizade², Enayatollah Bakhshi³

1. Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

2. Department of Radiology, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran.

3. Department of Biostatistics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.



Citation: Zarei Mahmood Abadi M, Yadegari F, Mehdizade M, Bakhshi E. Test-Retest and Inter-Rater Reliability Study of the Schedule for Oral-Motor Assessment in Persian Children. Iranian Rehabilitation Journal. 2018; 16(1):45-54. <https://doi.org/10.29252/NRIP.IRJ.16.1.45>

doi: <https://doi.org/10.29252/NRIP.IRJ.16.1.45>

Article info:

Received: 31 May 2017

Accepted: 14 Sep. 2017

Keywords:

Feeding behavior, Observation, Oral manifestation, Child

ABSTRACT

Objectives: Reliable and valid clinical tools to screen, diagnose, and describe eating functions and dysphagia in children are highly warranted. Today most specialists are aware of the role of assessment scales in the treatment of affected individuals. However, the problem is that the clinical tools used might be nonstandard, and worldwide, there is no integrated assessment performed to assess the aforementioned conditions. In a previous study, we translated and examined the schedule for oral-motor assessment for its content validity. Therefore, in this study, we aimed to investigate the test-retest and inter-rater reliability of the schedule for oral-motor assessment.

Methods: In this validation project, 23 children (6-48 month-old) who were referred to the imaging center were recruited. The assessment was based on the schedule for oral-motor assessment and was conducted by two speech-language pathologists for an inter-rater agreement evaluation. A retest was completed on 13 children up to 7 days. Intraclass Correlation Coefficient (ICC), kappa coefficient, and percentage of agreement were analyzed by using SPSS software (version 17).

Results: We found a moderate Intraclass Correlation Coefficient (ICC=0.48) and a relatively sensible kappa coefficient (K=0.50) for test-retest agreement. Two speech-language pathologists reached 91% agreement on normal-dysfunctional decisions of schedule for oral-motor assessment. Sections 2, 3, 4, and 6 of schedule for oral-motor assessment demonstrated an agreement of 100%, and Sections 1 and 5 demonstrated an agreement of 91%.

Discussion: Schedule for oral-motor assessment meets moderate levels of test-retest and inter-rater reliability criteria. Nonetheless, the percentage of decision consistency for inter-rater and test-retest examination turned out to be excellent.

* Corresponding Author:

Fariba Yadegari, PhD

Address: Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

Tel: +98 (21) 22180043

E-mail: faribayadegari@yahoo.com

1. Introduction

Development of new scales as assessment tools of swallowing and eating behaviors and dysphagia for children is increasing [1-4]. Nowadays, due to appropriate therapeutic techniques, more children and adults are being survived from medical accidents. Thus, swallowing disorders have higher rate of prevalence among children and adults particularly in disabled populations than that of normal adults [5-8].

Identification and describing dysphagia and feeding performance require standard screening and clinical assessments [9]. Leonard and Kenda noted the importance of clinical-observational assessments by which children at risk of feeding problems are identified. They considered this as essential not only in the therapeutic affairs but also in early intervention programs [10]. Currently, most specialists are aware about the roles of assessment scales in the assessment and remedy of the aforementioned disorders. The problem is that these tools are non-standard with sometimes undecided validity; there is no integrated assessment throughout the world [11].

Precise description of eating conducts in children requires development and improvement of the systems which assess complex set of interconnected motor skills based on observation. Such a system must be clinically valid and reliable [12]. One of the most famous methods of reliability assessment is test-retest in which a group is tested once and then is tested again in a specified time interval to calculate the correlation of the two tests [13]. Consequently, we can determine the reliability of oral-motor assessment scales with regard to the child's eating-related behaviors, for example, having the ability to hold the tongue up while being fed. The other method known is examining the correlation of two independent ratings of two different observers. This is called inter-rater reliability test. Kappa agreement coefficient is defined as an agreement index by Cohen which is used for reliability assessment [14]. We used this method to estimate the reliability of assessment scales of feeding behaviors in children, for example, how two raters score a child's normal and abnormal swallowing motor functions.

There are various scales in the assessment of pediatric sucking and swallowing; however, only a few of them have been validated, for example, Neonatal Oral-Motor Assessment Scale (NOMAS), developed by Palmer, is highly used in infant eating assessment. This is the only scale available to evaluate premature infants who are fed by breast and bottle [1, 9, 15], and it also assesses Nutri-

tive Sucking (NS) and Non-Nutritive Sucking (NNS) up to the age of 8 weeks. In order to assess the infant's sucking skills, the infant is observed for 2 min and information regarding his/her coordination between sucking and swallowing; maintaining the sucking, swallowing, and breathing coordination in bursts; and rhythmic sucking is collected [16, 17].

NOMAS with its 28 items classifies sucking patterns of the babies into 3 groups: 1. Normal sucking pattern which is shown by infants who can coordinate between sucking, swallowing, and breathing either in NS and NNS; 2. Disorganized sucking pattern that is observed in infants who are not able to coordinate swallowing and breathing actions, for example, infants suffering from breathing problems, a heart condition, or gastrointestinal problems; and 3. Dysfunctional sucking pattern which is observed in infants who have abnormal motions of jaws and tongue. It is also common in infants suffering from either neurological or anatomical disorder [12, 15].

Zarem et al. analyzed psychometric aspect of NOMAS in a prospective cohort performed on 75 premature infants (39 girls and 36 boys) born in 30th week of pregnancy or lower with average ages of 26.56 weeks and average weights of 967.33 g. They made a video of infants eating process before discharging from NICU. Accordingly, sucking pattern of the infant was classified as normal, disorganized, or dysfunctional. In order to assess the reliability, 6 raters selected 5 random films and defined infant's sucking pattern according to the scale. Then, 2 weeks later raters randomly observed the films again and defined the sucking pattern. This study proved that infants who showed the abnormal sucking pattern, based on NICU network neurobehavioral scale, confronted more stress. In addition, there was no significant correlation between NOMAS scores and results of children who were 2 years of age. This study showed that NOMAS could not predict developmental delays of at risk premature infants and the reliability of the finding were not ideal [15].

Costa and Schans studied the test-retest and inter-rater reliability of NOMAS on 75 premature infants (26-36 week). Four observers who were certified in performing the scale participated in their study. According to their results, test-retest reliability of NOMAS was found to be good or approximately excellent ($K=0.33-0.94$), whereas inter-rater reliability was found to be medium or excellent ($K=0.4-0.65$) [16]. Schedule for Oral-Motor Assessment (SOMA) is a scale of observing and recording the oral-motor skills of young children aged 8-24 months in order to identify their possible abnormalities related to

feeding. It assesses oral-motor abilities using foods with different textures at 3 levels: 1. Functional area such as lips, jaw, and tongue; 2. Functional units as the job of groups of muscles, for example, the role of lips in keeping food in the mouth; and 3. Discrete oral-motor behavior, for example, upper lip removes food from the spoon.

Reilly et al. studied test-retest reliability of SOMA. They studied typically developing children (n=58) with an average age of 12.2 months, Non Organic Failure to Thrive (NOFT) children with an average developmental age of 15 months (n=56), and children suffering from cerebral palsy with an average developmental age of 20.2 months (n=13). Test-retest agreement for 84% of the assessable behaviors proved to be outstanding ($\kappa=1$), and excellent inter-rater agreement was obtained for 62% of the distinct oral-motor behaviors ($K>0.75$) [12]. Skuse et al. studied the reliability of SOMA in a companion paper. As mentioned earlier, this scale was invented for the preverbal children with respect to identifying the defective eating-related organs. The research was administered on the same 127 children in 3 groups mentioned above. Their results showed that criterion validity was reasonably recognized by the analysis and also a shortened form of the SOMA -appropriate for screening- was introduced with a positive predictive validity larger than 90% and sensitivity higher than 85% useful in identifying infants with substantial oral-motor disabilities [18].

With respect to the importance of assessment of children's swallowing problems and the remedies which must be based on such assessment results, in this study, we aimed at answering the following questions: 1. Does SOMA contain good inter-rater reliability in Persian children? 2. Does SOMA contain good test-retest reliability in Persian children?

2. Methods

This validation study followed a previous study performed by Zarei and Yadegari on translation of Schedule for oral motor assessment and examining its content validity, Middle Eastern Journal of Disability Studies which introduced the Persian translated version of SOMA.

In this study, we used SOMA to identify dysfunctional areas that lead to eating disorders in children. It contains 65 items in 6 separate sections: each section examines one food texture. Scale sections are as follows: 1. Puree (yogurt) with spoon; 2. Semisolid (creamy cheese) with spoon; 3. Solid (cooked rice) with spoon; 4. Cracker (biscuit) through finger feeding; 5. Liquid (water/milk)

through breast feeding or bottle; and 6. Liquid (water) with cup. Scoring the scale comprises a "yes"/"no" system with a score of 0 given to "yes" and 1 for "no" to a normal function; a score of 1 given to "yes" and 0 for "no" to an abnormal function. Cutoff scores of the sections were as follows: section 1=3; section 2=4; section 3=4; section 4=9; section 5=5; and section 6=5.

Parents were asked to take their children to the speech therapy clinic of Children's Medical Center, Persia. A questionnaire with 3 sections was completed through interview with the parents: 1. Demographic characteristics; 2. Medical history; and 3. Feeding history. Inclusion criteria encompassed children of 6 to 48 months with a prescription of a barium swallow test who were sent to the radiology center of Children's Medical Center. Exclusion criteria were as follows: 1. 6 to 48 month children who could not be tested with barium swallow test; 2. Children who were under intubated feeding in the time of barium swallow test; and 3. Children with poor cooperation.

Consent forms were signed by main caregivers of the children (almost always their parents). If parents decided to leave the project, they were allowed to do so. No fee was charged for the experiments. The information of the participants was confidential. The research was assumed an ethical code (IR.USWR.REC.1395.316) by Ethics Board of the University of Social Welfare and Rehabilitation Sciences.

The procedure was explained by the first author to the parents mentioning five food textures (yoghurt, cream cheese, cooked rice, cracker, milk, and water regarding the habitual bottle or breast feeding of the child), observation process by Speech Language Pathologists (SLPs) and mentioned them that there were no side effects for children. As some children stopped eating due to the presence of the strangers, parents were instructed verbally and practically to feed children. Feeding scene was easily visible by the raters.

Parents were instructed to fill teaspoon by half while nourishing yoghurt, cream cheese, and cooked rice. They ought to use their fingers while feeding biscuit, keeping the thumb 3/4 inches away from the biscuit edge while holding their forefinger under the biscuit. Mothers were trained how to feed water by glass cups or how to milk by bottle or breast depending on the child's diet. They were taught not to transfer to any other food while using one and should wait for the rater's permission. SLP raters observed the process and completed SOMA items. If the observation was not satisfied for any reason, the mother was asked for a repetition. If the behavior was

seen, a “yes” was ticked on the form; otherwise, a “no” was selected. The assessment of five food textures was performed in triplicates [12, 18, 19].

In case that the child is inclined to test the food himself/herself, they were allowed to do so. In all stages, the child was sitting on the bed, but in the fifth stage, the child was in mother’s cuddle while being fed by breast or bottle. Food textures contained 100 g low fat pasteurized and homogenized yoghurt (with commercial name of Jal), pasteurized and homogenized cream cheese with 24% fat (with commercial name of Pegah), biscuit without trans fatty acid (with commercial name of Madar), cooked rice, mother’s or formula milk, and Tehran’s urban drinking water.

To evaluate the inter-rater reliability, 23 children who were referred for a barium swallow test were recruited. Children were assessed through SOMA by 2 SLPs with at least 8 years of experience with childhood dysphagia. In order to examine test-retest reliability, 13 children (5 girls and 8 boys) were assessed again within 7 days. Data were analyzed by SPSS (version 17). The Intraclass Correlation Coefficient (ICC), agreement percentage, and kappa agreement coefficient were calculated.

3. Results

Inter-rater agreement was completed within 5 months. The examined children were 10 girls and 13 boys in the range of 6 to 48 months of age (mean=18.13, SD=11.72) and birth weight in the range of 1500 g to 3900 g of them,

69.6% had a hospitalization history. Of them, 13% were found to experience breathing difficulty while feeding and 30.4% experienced nonoral feeding. Moreover, 34.8% of the children experienced tiredness while feeding and 34.8% used to spend more than half an hour eating a meal. In addition, 8.7% of the children suffered a growth delay and 60.9% were under anti-reflux medications.

Calculated ICC for inter-rater reliability was found to be 0.48. However, two SLPs had 91-100% agreement in their final decision of abnormality or normality of the functions through 6 stages of SOMA. The results indicated 100% agreement for sections 2 (semi-solid), 3 (solid), 4 (cracker), and 6 (liquid cup) and 91% agreement for sections 1 (puree) and 5 (liquid bottle). Calculated ICC for test-retest reliability was found to be 0.50. Table 1 contains the percentage of agreement and Kappa values of the part 1 of SOMA. As indicated in Table 1, all the items had above 76% agreement for test-retest and above 47% for inter-rater reliabilities. In addition, 87% of items revealed a Kappa coefficient between 0.23 and 0.68.

Table 2 contains the percentage of agreement and Kappa values of the part 2 of SOMA. As indicated in Table 2, all the items had above 69% agreement for test-retest and above 47% for inter-rater reliabilities. Furthermore, 62% of the items in part 2 obtained a Kappa value between 0.11 and 1.

Table 3 contains the percentage of agreement and Kappa values of the part 3 of SOMA. As indicated in Table 3,

Table 1. Test-retest and inter-rater reliability results of part 1 of SOMA

| Items of Part 1 SOMA (Puree) | | | Test-Retest Agreement Percent (%) | Test-Re-test Kappa Value | Inter-Rater Agreement Percent (%) | Inter-Rater Kappa Value |
|------------------------------|----------|---|-----------------------------------|--------------------------|-----------------------------------|-------------------------|
| 1 | React | Head orientation to spoon | 84 | 0.58 | 86 | 0.64 |
| 2 | Sequence | Smooth rhythmic sequence | 76 | -0.11 | 91 | 0.45 |
| 3 | Lip | Upper lip removes food from spoon | 84 | 0.58 | 86 | 0.68 |
| 4 | Lip | Lower/ upper lip assists in cleaning | 84 | 0a | 91 | 0.46 |
| 5 | Lip | Lower lip active during suck/ munch/ chew | 92 | 0a | 86 | 0.33 |
| 6 | Tongue | Consistent/ considerable protrusion | 92 | 0a | 86 | 0.33 |
| 7 | Tongue | Protrusion beyond incisors | 76 | -0.11 | 86 | 0.49 |
| 8 | Jaw | Graded jaw opening | 92 | 0a | 47 | -0.08 |

Iranian Rehabilitation Journal

0a: No statistics are computed because first rater response and second rater response are constants or because test response and retest response are constants.

Table 2. Test-retest and inter-rater reliability results of part 2 of SOMA

| Items of Part 2 Schedule for Oral Motor Assessment: (Semi-Solid) | | | Test-Retest Agreement Percentage (%) | Test-Retest Kappa Value | Inter-Rater Agreement Percentage (%) | Inter-Rater Kappa Value |
|--|------------|--|--------------------------------------|-------------------------|--------------------------------------|-------------------------|
| 1 | Drool | Consistent/ considerable drooling | 92 | 0a | 100 | 1 |
| 2 | Sequence | Smooth rhythmic sequence | 76 | -0.11 | 91 | 0.45 |
| 3 | Initiation | Sequence initiated within 2 seconds | 69 | 0a | 86 | 0a |
| 4 | Lip | Lip closed during opening | 84 | -0.08 | 47 | 0.11 |
| 5 | Jaw | Graded jaw opening | 100 | 0a | 47 | -0.08 |
| 6 | Jaw | Internal jaw stabilization | 84 | -0.08 | 100 | 0a |
| 7 | Jaw | External jaw stabilization required 100% | 76 | 0a | 100 | 1 |
| 8 | Jaw | Associated jaw movements | 76 | 0.67 | 100 | 1 |

Iranian Rehabilitation Journal

all the items had above 76% agreement for test-retest and above 43% for inter-rater reliabilities. Moreover, 66% of the items revealed a Kappa value between 0.18 and 0.86.

Table 4 contains the percentage of agreement and Kappa values of the part 4 of SOMA. As indicated in Table 4, all the items had above 53% agreement for test-retest and above 60% for inter-rater reliabilities. In addition, 90% of the items obtained a Kappa value between 0.04 and 0.83.

Table 5 contains the percentage of agreement and Kappa values of the part 5 and 6 of SOMA. As indicated in

Table 5, all the items had above 61% agreement for test-retest and above 47% for inter-rater reliabilities. In addition, 66% of the items in part 5 and 6 obtained a Kappa value between 0.08 and 1.

Test-retest reliability assessment took 4 months to complete; 13 children (5 girls, 8 boys) within the age range of 7-22 months (mean=12.07, SD=4.09) participated. Of them, 92.3% had a history of hospitalization, 23.1% suffered breathing distress while feeding, and 15.4% experienced nonoral feeding. In addition, 30.8% of the children experienced tiredness while eating and 38.5% used to spend more than half an hour for eating a meal.

Table 3. Test-retest and inter-rater reliability results of part 3 of SOMA

| Items of Part 3 Schedule for Oral Motor Assessment: (Solid) | | | Test-Retest Agreement Percentage (%) | Test-Retest Kappa Value | Inter-Rater Agreement Percentage (%) | Inter-Rater Kappa Value |
|---|-----------|--|--------------------------------------|-------------------------|--------------------------------------|-------------------------|
| 1 | Food loss | None/trivial | 76 | 0.67 | 73 | 0.46 |
| 2 | Drool | Consistent/ considerable drooling | 92 | 0.62 | 95 | 0.86 |
| 3 | Sequence | Smooth rhythmic sequence | 92 | 0.62 | 95 | 0a |
| 4 | Lip | Lower lip draws inwards around spoon | 100 | 1 | 91 | 0.45 |
| 5 | Lip | Upper lip removes food from spoon | 84 | 0.43 | 69 | 0.18 |
| 6 | Lip | Lower lip behind upper teeth/ sucking | 84 | 0.40 | 91 | 0.74 |
| 7 | Lip | Lower lip active during suck/ munch/chew | 92 | 0.62 | 100 | 0a |
| 8 | Tongue | Transient/ minimal tongue protrusion | 84 | 0.58 | 78 | 0.55 |
| 9 | Jaw | Graded jaw opening | 92 | 0.62 | 43 | -0.08 |

Iranian Rehabilitation Journal

Table 4. Test-retest and inter-rater reliability results of part 4 of SOMA

| | Items of Part 4 Schedule for Oral Motor Assessment (Cracker) | | Test-Retest Agreement Percentage (%) | Test-Retest Kappa Value | Inter-Rater Agreement Percentage (%) | Inter-Rater Kappa Value |
|----|--|--|--------------------------------------|-------------------------|--------------------------------------|-------------------------|
| 1 | Food loss | Profuse/ marked food loss | 92% | 0.75 | 86% | 0.68 |
| 2 | Drool | Profuse/ marked drooling | 92% | 0a | 95% | 0.83 |
| 3 | Initiation | Sequence initiated within 2 seconds | 92% | 0a | 82% | 0.24 |
| 4 | Lip | Lower lip behind upper teeth to suck | 92% | 0a | 65% | 0.24 |
| 5 | Lip | Lips close around stimulus during bite | 61% | 0.03 | 82% | 0.55 |
| 6 | Lip | Lips close intermittently during suck/ munch/ chew | 84% | 0a | 82% | 0.23 |
| 7 | Tongue | Transient/ minimal tongue protrusion | 61% | 0.31 | 65% | 0.31 |
| 8 | Tongue | Considerable/ consistent tongue protrusion | 92% | 0a | 95% | 0.77 |
| 9 | Tongue | Protrusion beyond incisors | 84% | 0a | 91% | 0.61 |
| 10 | Tongue | Protrusion beyond lips | 84% | -0.08 | 91% | 0.61 |
| 11 | Jaw | Internal jaw stabilization established | 92% | 0.62 | 78% | 0.16 |
| 12 | Jaw | Variable stabilization (not fully established) | 76% | 0.26 | 60% | 0.24 |
| 13 | Jaw | External stabilization | 84% | -0.08 | 69% | -0.14 |
| 14 | Jaw | Vertical movements | 92% | 0a | 100% | 0a |
| 15 | Jaw | Wide vertical excursions | 61% | 0.15 | 73% | 0.49 |
| 16 | Jaw | Small vertical excursions | 61% | 0.15 | 73% | 0.44 |
| 17 | Jaw | Associated head movements to bite | 84% | 0a | 95% | 0.77 |
| 18 | Jaw | Uses fingers to transfer food | 100% | 0a | 82% | 0.40 |
| 19 | Swallow | Gagging | 53% | 0.30- | 82% | 0.55 |
| 20 | Bite | Controlled sustained bite | 76% | 0.45 | 82% | 0.51 |
| 21 | Bite | Graded jaw opening | 100% | 0a | 82% | 0.04 |
| 22 | Bite | Mouths cracker only | 84% | 0.43 | 91% | 0.69 |

Iranian Rehabilitation Journal

4. Discussion

In Persia, as valid and reliable assessment scales of “swallowing” in children are scarce, we aimed to investigate the reliability of a previously translated scale called SOMA [20]. The purpose of this study was to examine the test-retest and inter-rater reliability of the Persian version of SOMA.

According to Reilly et al.’s study, test-retest reliability showed that 84% of the ratings revealed an

outstanding agreement (Kappa=1), 6.2% were larger than 0.75 which was considered an excellent agreement, 8.7% were between 0.40 and 0.75, and 1.1% were lower than 0.40 [12]. If we put our own results of test-retest reliability into a standard interpretation, the results showed that 10.34% of discrete oral-motor behaviors were of an excellent agreement (K=1), 31.03% of discrete oral-motor behaviors were of significant agreement (K=0.61-0.80), 31.03% of them proved to have an average agreement (K=0.41-0.60), 17.24% of discrete oral-motor behaviors were of poor agreement

Table 5. Test-retest and inter-rater reliability results of part 5 and 6 of SOMA

| Items of Part 5 and 6 Schedule for Oral Motor Assessment (Liquid Bottle/Liquid-Cup) | | | Test-Retest Agreement Percentage(%) | Test-Re-test Kappa Value | Inter-Rater Agreement Percentage(%) | Inter-Rater Kappa Value |
|---|-------------|--|-------------------------------------|--------------------------|-------------------------------------|-------------------------|
| 1 | React | Anticipatory mouth opening | 84% | 0.08- | 100% | 1 |
| 2 | React | No liquid enters mouth | 84% | 0.43 | 95% | 0.91 |
| 3 | Accept | Accepts liquid within 2 seconds | 76% | 0.31 | 95% | 0.64 |
| 4 | Lip | Upper lip firmly seals around teat | 100% | 1 | 95% | 0.77 |
| 5 | Lip | Intermittent/ incomplete upper lip contact/seal | 84% | 0.58 | 91% | 0.82 |
| 6 | Lip | Intermittent/ incomplete lower lip contact/ seal | 84% | 0.58 | 47% | 0.64 |
| 7 | Lip | Lip closure during swallow | 84% | 0a | 100% | 1 |
| 8 | Jaw | Small vertical movements | 69% | 0a | 52% | 0.08 |
| 9 | Sequence | Smooth rhythmic sequence | 92% | 0a | 95% | 0.64 |
| 10 | Accept | Accepts within 2 seconds | 100% | 1 | 95% | 0.77 |
| 11 | Sequencing | Panic reactions when liquid placed in mouth | 92% | 0a | 95% | 0a |
| 12 | Sequencing | Choking | 92% | 0a | 100% | 0a |
| 13 | Liquid loss | Profuse/ marked liquid loss | 69% | 0.35 | 82% | 0.64 |
| 14 | Tongue | Tongue thrust | 92% | 0.62 | 86% | 0.65 |
| 15 | Tongue | Asymmetry | 92% | 0a | 86% | -0.06 |
| 16 | Jaw | Small vertical movements | 84% | -0.08 | 73% | -0.15 |
| 17 | Jaw | Jaw clenching | 84% | 0a | 82% | -0.09 |
| 18 | Swallow | Gagging | 61% | -0.22 | 86% | -0.06 |

Iranian Rehabilitation Journal

($K=0.21-0.40$) and 10.36% were of very poor agreement ($K<0.21$) [14]. These results came from those items that their Kappa value was calculable.

In addition, Reilly et al. analyzed an inter-rater agreement of pass/fail of the items that showed the following results. Furthermore, 68% of evaluations led to perfect agreement, 9% were larger than 0.75, 13% of the ratings were between 0.40 and 0.75, and 2% were lower than 0.40 [12]. Inter-rater reliability results of this study showed that 18% of discrete oral-motor behaviors had total agreement ($K=1$), 30% of the discrete oral-motor behaviors revealed significant agreement ($K=0.80-0.61$), 24% of the behaviors proved to have an average agreement ($K=0.60-0.41$), 12% of the discrete oral-motor behaviors had lower-average agreement ($K=0.41-0.21$),

and 10.36% of the behaviors were of very poor agreement ($K<0.21$).

Cořta and Schans (2008) analyzed the reliability of Neonatal Oral-Motor Assessment Scale (NOMAS). According to their findings, NOMAS proved to have reasonable to excellent test-retest agreement ($K=0.33-0.94$), whereas inter-rater agreement was found to be modest to considerable ($K=0.40-0.65$) [16]. Therefore, the test-retest agreement of NOMAS in their study was higher than inter-rater agreement, a result which is contrasting to our results, that is, test-retest agreement was lower than inter-rater agreement. Only 32.28% of the items of the Persian version of SOMA had average and higher than average test-retest agreement, whereas 56.91% of the items showed average and higher than average.

Skuse et al. (1995) introduced methods of validation that showed SOMA analysis of test-retest and inter-rater reliability was excellent [18], which is slightly not consistent with our findings. It is noteworthy that analysis in some items of the SOMA is easier than the others. For example, the items which are concerned to oral- motor functions of the child rather than discrete oral-motor behaviors, obtaining a good agreement among raters may be hard. Looking practically, when SLPs were asked to judge an oral-motor functional unit such as “biting” skill as a whole, they were not in a total agreement, but when asked if there is a controlled stable biting, they may well agree [12].

Our study was limited due to not supplying videos, rather we administered live observations. SOMA assessment is better run when there are videos to be observed repeatedly, yet this way is not perfect either. This may be owing to the intrinsic characteristics of such observations; they seem dependent to the individual rater experience; these observations seem somewhat confusing and unstable for rating. Finally, we suggest that in order to get higher kappa values and agreement percentage, we should enhance the number of the participants.

5. Conclusion

Persian version of SOMA proved to have a very good inter-rater and test-retest agreement percentage. Inter-rater agreement percentage was excellent (91%) and in test-retest agreement percent good (84%). In addition, 57% of the items in SOMA for inter-rater reliability were shown to be average and above and 33% of the items for test-retest reliability were shown to be average and above.

Acknowledgements

This paper was extracted from the MSc. thesis of the first author in the Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. The authors wish to thank parents who cooperated in this study. They also wish to thank Mrs. Maryam Khodayar who cooperated actively as an speech-language pathologists rater.

Conflict of Interest

The authors declared no conflicts of interest.

References

- [1] Rogers B, Arvedson J. Assessment of infant oral sensorimotor and swallowing function. *Mental Retardation and Developmental Disabilities Research Reviews*. 2005; 11(1):74-82. doi: 10.1002/mrdd.20055
- [2] Thoyre SM, Pados BF, Park J, Estrem H, Hodges EA, McComish C, et al. Development and content validation of the pediatric eating assessment tool (Pedi-EAT). *American Journal of Speech-Language Pathology*. 2014; 23(1):46. doi: 10.1044/1058-0360(2013/12-0069)
- [3] Allen SL, Smith IM, Duku E, Vaillancourt T, Szatmari P, Bryson S, et al. Behavioral pediatrics feeding assessment scale in young children with autism spectrum disorder: Psychometrics and associations with child and parent variables. *Journal of Pediatric Psychology*. 2015; 40(6):581-90. doi: 10.1093/jpepsy/jsv006
- [4] Serel Arslan S, Demir N, Karaduman AA, Belafsky PC. The pediatric version of the eating assessment tool: A caregiver administered dysphagia-specific outcome instrument for children. *Disability and Rehabilitation*. 2017; 1-5. doi: 10.1080/09638288.2017.1323235
- [5] Romano M, Schultz T, Tai A. The diagnostic test accuracy of clinical swallow assessment for oropharyngeal aspiration: A systematic review. *JBIC Database of Systematic Reviews and Implementation Reports*. 2012; 10(56):1-16. doi: 10.11124/01938924-201210561-00002
- [6] Kakodkar K, Schroeder JW. Pediatric dysphagia. *Pediatric Clinics of North America*. 2013; 60(4):969-77. doi: 10.1016/j.pcl.2013.04.010
- [7] Alberta Health Services. Oral feeding reference poster [Internet]. 2016 [Updated 2016 April 20]. Available from: <https://www.albertahealthservices.ca/assets/programs/ps-1055678-neonatal-oral-feeding-reference-poster.pdf>
- [8] Heckathorn D-E, Speyer R, Taylor J, Cordier R. Systematic review: Non-instrumental swallowing and feeding assessments in pediatrics. *Dysphagia*. 2015; 31(1):1-23. doi: 10.1007/s00455-015-9667-5
- [9] Arvedson JC. Assessment of pediatric dysphagia and feeding disorders: Clinical and instrumental approaches. *Developmental Disabilities Research Reviews*. 2008; 14(2):118-27. doi: 10.1002/drr.17
- [10] Lucas G. Critical review: Exploring the inter-rater reliability of two assessment tools used to identify feeding problems in neonates. London: University of Western Ontario; 2009.
- [11] Arvedson JC, Brodsky L. *Pediatric swallowing and feeding: Assessment and management*. Boston: Cengage Learning; 2002.
- [12] Reilly S, Skuse D, Mathisen B, Wolke D. The objective rating of oral-motor functions during feeding. *Dysphagia*. 1995; 10(3):177-91. doi: 10.1007/bf00260975
- [13] Seif A. [Measurement, assessment and evaluation of education (Persian)]. Tehran: Dowran; 2016.
- [14] Pashashrifi H, Sharifi N. [Principles of psychometrics and psychological testing (Persian)]. Tehran: Roshd; 2014.
- [15] Zarem C, Kidokoro H, Neil J, Wallendorf M, Inder T, Pine-da R. Psychometrics of the neonatal oral motor assessment

- scale. *Developmental Medicine & Child Neurology*. 2013; 55(12):1115-20. doi: 10.1111/dmcn.12202
- [16] Da Costa SP. [Development of sucking patterns in preterm infants [PhD thesis]. Groningen: University Library Groningen; 2010.
- [17] Bingham PM, Ashikaga T, Abbasi S. Relationship of neonatal oral motor assessment scale to feeding performance of premature infants. *Journal of Neonatal Nursing*. 2012; 18(1):30-6. doi: 10.1016/j.jnn.2010.09.004
- [18] Skuse D, Stevenson J, Reilly S, Mathisen B. Schedule for oral-motor assessment (SOMA): Methods of validation. *Dysphagia*. 1995; 10(3):192-202. doi: 10.1007/bf00260976
- [19] Ko MJ, Kang MJ, Ko KJ, Ki YO, Chang HJ, Kwon JY. Clinical usefulness of schedule for oral-motor assessment (SOMA) in children with dysphagia. *Annals of Rehabilitation Medicine*. 2011; 35(4):477. doi: 10.5535/arm.2011.35.4.477
- [20] Zarei Mahmood Abadi M, Yadegari F, Bakhshi E. Introduction of schedule for oral-motor assessment in Persian and examining its validity. *Middle Eastern Journal of Disability Studies*. In Press.

