IZMIR DEMOCRACY UNIVERSITY HEALTH SCIENCES JOURNAL
IDUHES
e-ISNN:2651-4575

IDUHeS, 2020; 3(2): 135-144

Araştırma Makalesi– Research Paper

6 VE 16 YAŞ OKUL ÇOCUKLARI ARASINDAKİ SSI-4-TR / KEŞİDA-4 DÖRDÜNCÜ BASKISININ TÜRKÇE VERSİYONUNUN GÜVENİLİRLİĞİ VE GEÇERLİLİĞİ

RELIABILITY AND VALIDITY OF THE TURKISH VERSION OF THE STUTTERING SEVERITY INSTRUMENT, FOURTH EDITION (SSI-4-TR/ KEŞİDA-4) BETWEEN 6 AND 16 YEARS OLD SCHOOL CHILDREN

Ayşe İlayda MUTLU[1], Şadiye BACIK TIRANK[2], Bülent GÜNDÜZ[2]

Özet

Kekemeliğin şiddetini belirlemek ve tedavi sonuçlarını değerlendirmek için standart ve güvenilir bir araca ihtiyaç bulunmaktadır. Kekemeliği ölçmek için en güvenilir araçlardan biri Kekemelik Önem Düzeyi - Dördüncü Baskısı'dır (SSI-4). Bu çalışmanın amacı, okul çağındaki kekemeli çocuklarda Kekemelik Şiddet Ölçeği'nin (SSI-4-TR) Türkçe versiyonunun geçerlilik ve güvenilirliğini değerlendirmektir. Başvuru süreci için orijinal SSI-4 alınmıştır. Türkçe uygulamasının çevirisi üç bağımsız uzman tarafından değerlendirildikten sonra 6-16 yaş arası 47 çocuk çalışmaya dahil edilmiş ve anketleri tamamlanmıştır. Kekemelik engeli olan çocuklar, SSI-4 ile test edilmiş ve kekemeliğin şiddeti uygulayıcı tarafından kaydedilmiştir. Veriler, test-tekrar test yöntemine dahil edilen rastgele seçilen 18 çocuktan toplanmıştır. Uygulayıcıya ek olarak üç bağımsız uzman tarafından değerlendirilmiştir. SSI-4'ün en yüksek güvenilirlik ortalaması 0,94 ve maksimum mantıksal doğrulama istatistiği 0,846 olarak kaydedilmiştir. Kekemelik şiddeti değerlendirmesi için 0.73 ile 0.903 arasındaki bir ritim boşluğu optimum olarak kabul edilmiştir. Türkiye SGK-4'ün genel güvenilirlik düzeyi için, Cronbach's alpha, güvenilir bir kekemelik değerlendirme aracı olduğunu gösteren 0.94 olarak kaydedilmiştir. Test-tekrar test yöntemi, gözlemcilerin güvenilirlik çalışmaları ve iç tutarlılık oranı, SSI-4'ün oldukça güvenilir olduğu kanıtlanmıştır. SGK-4'ün tercüme edilmiş Türkçe versiyonu kekemelik değerlendirmesi için güvenilir ve geçerli bir araçtır.

Anahtar kelimeler: SSI-4, SSI-4-TR, kekemelik şiddeti, kekemelik şiddeti, kekemelik için değerlendirme cihazı, KEŞİDA-4

Abstract

In order to determine the severity of stuttering and to evaluate the results of treatment, we need a standard and reliable instrument. One of the most reliable tools for measuring stuttering is the Stuttering Severity Instrument – Fourth Edition (SSI-4). The purpose of this study was to evaluate intra and inter-rater reliability of the Turkish version of the Stuttering Severity Instrument (SSI- 4-TR) in school-aged children with stuttering. For the application process, the original SSI-4 was obtained. After the translation for the Turkish application had been completed by three independent experts, 47 children aged between 6 and 16 were included in the study, and their questionnaires were completed. Children with a stuttering impediment were tested with the SSI-4, and the severity of stuttering was recorded by the practitioner. The data were collected from 18 randomly chosen children who were included in the test-retest method. They were evaluated by three independent experts in addition to the practitioner. The highest reliability average of SSI-4 was recorded as 0.94 and the maximum statistic of logical validation as 0.846. A rhythm gap from 0.73 to 0.903 is accepted as optimum for stuttering severity assessment. For the overall reliability level of the Turkish SSI-4, Cronbach's alpha was recorded as 0.94, showing that it is a credible stuttering assessment device. The test-retest method, credibility studies of observers, and the internal consistency ratio proved that the SSI-4 is highly reliable. The translated Turkish version of SSI-4 is a reliable and valid instrument for stuttering evaluation.

Key words: SSI-4, SSI-4-TR, assessment device for stuttering severity, stuttering severity, stuttering, KEŞİDA-4

Geliş Tarihi (Received Date): 12.05.2020, Kabul Tarihi (Accepted Date): 26.08.2020, Basım Tarihi (Published Date): 30.09.2020
[1]Özel Eğitim ve Danışmanlık Merkezi, Çankaya, Ankara, Türkiye, [2]Gazi Üniversitesi Sağlık Bilimleri Fakültesi Dil ve Konuşma Terapisi Bölümü, Ankara, Türkiye, E-mail:, sadiyebacik@gmail.com, **ORCID ID's**; Ş.B.T; https://orcid.org/0000-0002-3378-0862, A.İ.M.; https://orcid.org/0000-0001-6662-1553, B.G.; https://orcid.org/0000-0001-9826-7990

1. INTRODUCTION

Stuttering is a communication disorder which emerges as a result of the prevention of fluent speech due to repetitions, pauses, prolonging, and the gestures and mimicking that accompany them. Some changes in the speed, rhythm, emphasis, volume, and articulation of speech, as well as in the facial expressions and posture of the individual, can also be observed during speech (ASHA, 1993).

Although stuttering can be seen in all age groups, the problem is most frequently encountered during childhood. Most of the cases appear during the first phases of the development of language and speech, the preschool period. When children begin to form simple sentences with more than one word, some may start to stutter, and approximately 8.5% of 3-year-olds are diagnosed with stuttering. Among these 3-year-olds, 70% to 80% will resolve their stuttering without any direct intervention. However, the remaining 20% to 30% may continue to stutter in their school years and have a permanent problem [ASHA, 1993; Boey et al., 2007, pp. 310-329; Conture, 2001; Fibiger, 2010; George and Mallery, 2003).

There is surely a need for standard test devices to define, classify, and follow up on stuttering. In particular, there is an increasing requirement for devices that allow follow-up on therapeutic intervention and evaluate the speech of individuals in terms of time. Various assessment tools have been developed to meet these needs, particularly as to stuttering in children [Gilliam, 2009; Haynes and Pindzola 1997, pp. 238-245; Nippold, 2018, pp.4-12).

The test devices used in addition to natural observation analyze the different dimensions of the problem. The type of non-fluency is evaluated with the Dysfluency Type Index (DTI). The frequency and duration of stuttering is measured by the percentage of syllables stuttered per number of iterations (SS%), and the speech rate is measured by syllables per minute (SPM) and words per minute (WPM). The fluency and nature of speech is evaluated with the Naturalness Rating Scale (NAT), childhood stuttering with the Test of Childhood Stuttering (TOCS), and the physical aspects that accompany speech with Riley's Stuttering Severity Instrument, Fourth Edition (SSI-4) (Peters and Guitar, 1991).

The SSI-4 is widely used in the assessment of the severity of stuttering. It is a tool which can be used to test children, adults, literate individuals, or illiterate individuals. The test evaluates the naturality of speech by reading or speaking on any given subject, the severity of stuttering by calculating the number of syllables stuttered and the total number of syllables uttered by the speaker, the duration of stuttering by the longest case of stuttering, and the combined motor movements by the physical behaviors that accompany stuttering. Stuttering can be defined, using the total score attained through observation of stuttering frequency (%SS), average duration of the three longest stuttering moments, and the total scale values of physical concomitant behaviours. These are combined into one stuttering severity score (Riley and Bakker, 2009).

In Turkish literature, since there is no standard evaluation scale on stuttering, many problems are encountered in the definition, evaluation, and follow-up for stuttering. Culturally adapting SSI-4 to Turkish not only provides clinicians with the ability to conduct cross-cultural studies but also provides information about the influence of language on continuous speech assessment. Therefore, the objective of the current research was to translate the SSI-4 into

Turkish for school children aged 6–16. The Computerized Scoring of Stuttering Severity software (CSSS-2, version 2) was used in the present study to facilitate the calculation of stuttering frequency and duration.

2. MATERIALS AND METHODS

This study, following approval from the Gazi University Research Ethics Board (date: 24.03.2014; number: 25901600-1734), was conducted at Gazi University.

2.1. Translation Process

Permission was obtained from PRO-ED Inc., which has broadcast rights, on April 3, 2014. The team, including one linguist (an instructor at the Western Languages and Literature English Language and Literature Department) and one SLT (an instructor at the Prof. Dr. Necmettin Akyıldız Hearing, Speech, Voice, and Balance Disorder Center) and a sworn translator developed the SSI-4 sentences by following the phonetic rules described in the SSI-4 protocol (12). The team translated the SSI-4, and the translations were then compared and controlled by a specialist experienced in stuttering. Afterward, the Turkish version of SSI-4 was translated into English again and compared with the original. As a result, the original SSI-4 has been completely translated into Turkish without changing any of its parts or its meaning.

2.2. Participants

The study participants were recruited from patients who visited the Prof. Dr. Necmettin Akyıldız Hearing Speech Voice and Balance Disorder Center of the Gazi University Faculty of Medicine, Department of Otolaryngology & Audiology. All the evaluations were carried out in the Ear, Nose, and Throat Department and the Speech Language Pathology Department at Gazi University Hospital. Forty-seven children aged six through sixteen, who nominated themselves as stutterers and also stuttered in the preliminary informal interview with a SLP, were included in the study. Eighteen of the individuals participated in the test-retest activity. The following were the participation criteria for the study: (a) being aged 6–16, (b) being literate, (c) being stutterers, (d) not having a diagnosed or reported hearing problem, (e) being in conformity with their peers mentally and physically, and (f) not having organic problems in the orofacial area.

2.3. Judges

The judges were three SLPs who had at least three years' experience in the assessment and treatment of stuttering. The task for the two judges was to measure the stuttering severity of all 47 participants in this study using the Turkish version of the SSI-4. They received two training sessions in which they were familiarized with the Turkish version of the SSI-4 and the CSSS-2. We emphasize that they were experienced judges with extensive knowledge and experience in assessing and treating stuttering.

2.4. Procedure

First, a personal information form was filled out for the participating children. Second, the SSI-4 was applied through viewing a video recording of the children's speech and reading samples. Through the SSI-4, the frequency of stuttering, the duration of stuttering, the physical results which appeared as a result of stuttering, and the naturality of the individual's speech were evaluated.

After the non-fluency types of all children was determined by listening to their speech recordings, the practitioner watched the video recordings of all children who participated in the study once again after an interval of one month to determine the reliability of this data and to be able to calculate the reliability of intra-observers and interobservers. The frequency of the non-fluency types were rerecorded. The other observers watched the video recordings of the participating children and determined the non-fluency types and the secondary behaviors independently.

In the test-retest study, the same 18 individuals were tested again after one week following the first application. The consistency of the test scores was analyzed. The scores of the individuals obtained in the first evaluation were used in the adaptation study of SSI-4 to Turkish.

2.5. Application of the Test

First, a personal information form was filled out for the participating children. Second, the SSI-4 was applied through viewing a video recording of the children's speech and reading samples. Through the SSI-4, the frequency of stuttering, the duration of stuttering, the physical results which appeared as a result of stuttering, and the naturality of the individual's speech were evaluated.

After the non-fluency types of all children was determined by listening to their speech recordings, the practitioner watched the video recordings of all children who participated in the study once again after an interval of one month to determine the reliability of this data and to be able to calculate the reliability of intra-observers and interobservers. The frequency of the non-fluency types were rerecorded. The other observers watched the video recordings of the participating children and determined the non-fluency types and the secondary behaviors independently.

In the test-retest study, the same 18 individuals were tested again after one week following the first application. The consistency of the test scores was analyzed. The scores of the individuals obtained in the first evaluation were used in the adaptation study of SSI-4 to Turkish.

2.5.1. Frequency Score

Video and audio recordings were made of the reading assignment and the speaking samples. The reading material consisted of around 150–300 syllables, suitable to the age of the participants, and the speaking samples consisted of natural reciprocal speech on familiar subjects, again suitable to the participants' age.

The CSSS-2.0 software (version 2) was used to calculate the frequency and duration of stuttering, as well as to determine the stuttering instances during the reading assignment and the three longest stuttering instances on the part of the individuals. CSSS-2.0 requires clicking the left mouse button for each fluent syllable and holding down the right mouse button for the duration of the syllable for each stuttering. This method automatically produces the records of the percentage (frequency) and duration of the longest three instances of stuttered syllables.

Separate calculations for the frequency of stuttering were made for the reading and speaking assignments. The value was used to determine the assignment score through the table

in the researcher record form. These two assignment scores were added to give a result between 0 and 18, and the frequency score was obtained.

2.5.2. Duration Score

While the duration of stuttering was being calculated, the duration of the three longest stuttering instances were measured in seconds with a chronometer. These three durations were added to each other, and the average duration was obtained by dividing this number by three. Since the measurement of times shorter than a second is difficult, these were accepted as "very short" or "half a second" and were recorded respectively as two or four pause scores. After the average score was calculated, it was changed into a scale score.

2.5.3. Physical Results

The physical evaluation of the speech samples were scored as follows: 0 (no physical results), 1 (not recognizable unless viewed), 2 (a normal observer might perceive it slightly), 3 (distracting), 4 (very distracting), and 5 (severe and painful-looking).

2.5.4. Total Score

The total score was obtained by adding the frequency, duration and physical result scores to each other. The relative severity of stuttering for the individual in question was obtained by transforming the total score to percentage value or severity equivalent (Table 1).

Table 1. Total Scores For School Children

Total Score	Percentage	Severity
6–8	1–4	Mild
9–11	5–11	
11–15	12–23	Moderate
16–20	24–40	
21–23	41–60	Moderately severe
24–27	61–77	
28–31	78–88	Severe
32–35	89–95	
36 and over	96–99	Profound

2.5.5. Naturality

After the scoring of the speech samples, the naturality of speech was analyzed. The naturality of speech has been evaluated through the visual analogue scale from 1 (very natural speech) to 9 (very unnatural speech).

2.6. Statistical Characteristics

Statistical tests were performed using Statistical Package for the Social Sciences (SPSS) software, version 18 (SPSS, Inc., Chicago, IL). The Mann-Whitney U test and Kruskal-Wallis test were used to determine differences in nonparametric data. A one-way ANOVA analysis was used to determine data showing normal distribution. The total score and sub-scores of test-retest values were analyzed with the t-test in dependent groups. The relationships between the groups were analyzed by obtaining the Pearson correlation coefficient. The reliability of the tool was determined by Cronbach's alpha.

3.RESULTS

All 47 participants completed the SSI-4 tasks, and their speech and reading samples were assessed by three experienced judges.

Table 2. Demographic Findings

		N	%
Sex	Male	38	80.9
	Female	9	19.1
Stuttering severity	Mild	3	6.4
	Moderate	14	29.8
	Moderately severe	15	31.9
	Severe	11	23.4
	Profound	4	8.5
Naturality of speech	1	29	61.7
	2	5	10.6
	3	9	19.1
	4	0	0
	5	3	6.4
	6	1	2.1

The values related to the speech samples obtained from the patients are presented in Table 3.

Table 3. Findings Related To The Speech Samples Of The Patients

	Minimum	Maksimum	Mean ± Standard Dev
Number of syllables in the speech	172	500	399.9 ± 84.5
Number of stutters in the speech	5	82	38.1 ± 20.4
% of SS in the speaking assignment	1	26	9.7 ± 6.2
Speech score	2	9	6.4 ± 1.7
Frequency score	4	17	12.6 ± 3.1
Duration of stuttering	0	8	2.0 ± 1.4
Duration score	2	12	6.9 ± 2.7

Secondary behaviors	1	12	7.3 ± 2.6
Secondary behavior score	1	12	3.9 ± 2.6
Total score	7	38	23.4 ± 7.4
Severity of stuttering	1	5	3.0 ± 1.1
Naturality of speech	1	6	1.9 ± 1.3
Problem appearing age	2	12	4.7 ± 2.0

3.1. Reliability Of The Turkish Version Of The SSI-4

3.1.1. Inter-Judge Reliability

In the inter-judge evaluation, a meaningful difference was not observed in the total score ($p = 0.993$, $p > 0.05$), frequency score ($p = 0.949$, $p > 0.05$), reading score ($p = 0.956$, $p > 0.05$), speech score ($p = 0.981$, $p > 0.05$), duration score ($p = 0.998$, $p > 0.05$), or secondary behavior score ($p = 0.999$, $p > 0.05$).

3.1.2. Test-Retest Reliability

The test has been applied to the same patient group twice and it has been analyzed whether there is a meaningful difference between the two tests. A meaningful difference was found in the total score of the test ($p = 0.875$, $p > 0.05$) and its whole, except for the secondary behavior score, which is one of the sub-dimensions which constitutes the total ($p > 0.05$). Only the secondary behavior score displayed a meaningful difference ($p = 0.029$, $p > 0.05$).

The relationship between the SSI-4-TR total score and the frequency score is $r = 0.877$. The total score's relationship with the speech score is $r = 0.847$; with the reading score, $r = 0.730$; with the duration score, $r = 0.903$, and with the secondary behavior score, $r = 0.840$. The relationship between all these consistencies and between the subcomponents is statistically meaningful at the $p > 0.01$ level. Of all studies on the SSI-4, the highest reliability average recorded for the SSI-4-TR is 0.94, and the highest statistical logical validation is 0.846. The Cronbach's alpha has been determined to be 0.94, so the test can be said to have a high level of reliability.

4. DISCUSSION

The SSI-4 is a norm-referenced evaluation for stuttering, and its validity and reliability have been proven. There are no evaluation tools whose Turkish adaptations have been completed or are currently under development in our country for this purpose. For this reason, this study has aimed at creating a Turkish adaptation of the SSI-4.

Its reliability for the frequency score was determined as an average of 97.8%, the duration score as an average of 98.5%, the secondary behavior score as an average of 94.65%, and the total score as an average of 98.5% for interobservers. In the development of Riley's original SSI-4 (Riley, 2009), the reliability for the frequency score was determined as an average of 91%, the duration score as an average of 87.8%, the secondary behavior score as an average of 82.9% and the total score as an average of 93.4% for interobservers. Zolfaghari et al. (2014) found interobserver reliability of between 94% and 98% in their study. Our findings are higher than those in both studies in respect to the reliability for interobservers.

The total score obtained in the two different evaluations carried out in the SSI-4's test-retest reliability analysis (p = 0.875), the frequency score (p = 0.286), the reading score (p = 0.090), the speech score (p = 1.000), the duration score (p = 0.381) and the secondary behavior score (p = 0.029) were evaluated by a one-way ANOVA. A statistically meaningful difference between the observers was not found, with the exception of the secondary behavior score (p > 0.05). This difference which has emerged in the secondary behavior score is believed to result from the evaluation of the secondary behavior scores directly between the 0–5 interval, contrary to the other sub-calculations and the total score and its evaluation on the basis of a Likert-type scale. This displays the effectiveness of personal comments on the evaluation.

When the SSI-4-TR internal consistency was evaluated, we obtained reliability at a Cronbach's alpha level of 0.94. George and Mallery (Yairi and Seery, 2015) have given the following standards for the values of Cronbach's alpha: >.9 is considered to be perfect, >.8 is considered good, >.7 is considered acceptable, >.6 is considered questionable, >.5 is considered weak, and <.5 is considered unacceptable (Packman et al., 2007, p. 20).

When evaluated in this light, the 0.94 Cronbach's alpha value is at a perfect level and points out to the high level of our study's internal consistency. In terms of the internal consistency level, Zolfaghari et al. (2014) found the Cronbach's alpha level to be 0.87. The results obtained in this study were higher than those obtained in 2014.

As Riley (2009) has also stated, there are two basic difficulties in the reliability studies of tools used to measure the severity of stuttering: (1) there is no definite consensus between researchers and clinicians on the definition of stuttering, and (2) internal processes (such as embarrassment or avoidance) that are a serious component of stuttering are difficult factors for the individual with a stuttering problem to recognize and to evaluate for the observers. Thus, we have analyzed the relationship between the frequency of stuttering, which is the most frequently used behavioral measurement of stuttering, and the total score as Riley and have obtained a value of r = .877. This value is statistically meaningful at the level of p < 0.01. Riley (2009) obtained a value of r = .795 in schoolchildren.

When considered in this light, our Turkish adaptation's criteria validity is high and reliable. Our study has determined that the relationship between all the sub-components is statistically meaningful at the p < .01 level. Riley (2009) has also made the same evaluation in the development process of SSI-4-TR. The values we have obtained in our study are very similar to Riley's and they point to an even higher relationship. In this respect, the SSI-4-TR is a tool with a high structural validity.

Future studies should be aimed at creating a Turkish adaptation of SSI-4 for children who experience stuttering problems in the preschool period and for adults who stutter. Standardization studies should be carried out on the SSI-4 that has now been adapted to Turkish for school children aged 6–16.

The authors would like to thank Prof. Dr. Maviş Emel Kulak Kayıkcı for their thoughtful suggestions regarding both the development of this project and revision of this manuscript. In addition, the authors would like to thank the participants and their families for their significant contribution to this study.

Research and Publication Ethics Statement

This study meets the principles of research ethics. It is a part of a MSc study so it is ensured that it addresses ethics principles and standards of scientific research.

5. REFERENCES

ASHA: definitions of communication disorders and variations [Internet]. American Speech-Language-Hearing Association; 1993. Available from: www.asha.org/policy

Boey RA, Wuyts FL, Van de Heyning PH, De Bodt MS, et al. Characteristics of stuttering-like disfluencies in Dutch-speaking children. Journal of Fluency Disorders 2007;32(4):310–329.

Conture EG. Stuttering: its nature, diagnosis, and treatment. Boston (MA): Allyn & Bacon; 2001.

Fibiger S. Stuttering. In: International encyclopedia of rehabilitation. 2010.

George D, Mallery P. SPSS for Windows step by step: a sample guide and reference. Boston (MA): Allyn & Bacon; 2003.

Gillam RB, Logan KJ, Pearson NA. TOCS: test of childhood stuttering. Austin (TX): PRO-ED; 2009.

Haynes WO, Pindzola RH. Diagnosis and evaluation in speech pathology. 5th ed. Boston (MA): Allyn & Bacon; 1997. p. 238–245.

Nippold MA. Stuttering in preschool children: direct versus indirect treatment. Language, Speech, and Hearing Services in Schools 2018; 49(1): 4–12.

Packman A, Code C, Onslow M. On the cause of stuttering: integrating theory with brain and behavioral research. Journal of Neurolinguistics 2007; 20.

Peters TJ, Guitar B. Stuttering: an integrated approach to its nature and treatment. 1991.

Riley GD, Bakker K. Stuttering Severity Instrument: SSI-4. Austin (TX): PRO-ED; 2009.

Riley, GD. Stuttering Severity Instrument for children and adults (SSI-4). 4th ed. Austin (TX): PRO-ED; 2009.

Yairi E, Ambrose N. Early childhood stuttering. Austin (TX): PRO-ED; 2005.

Yairi E, Seery CH. Stuttering: foundations and clinical applications. Boston (MA): Pearson; 2015.

Zolfaghari M, Shafiei B, Tahmasebi Garmatani N, et al. Reliability of the Persian version of the Stuttering Severity Instrument (SSI-4) for preschool-age children. Middle Eastern Journal of Disability Studies 2014; 4(2): 20–25.