Development, Validity, and Reliability of the Auditory and Speech Performance Test for Children

Développement, validité et fiabilité du *Auditory and Speech Performance Test for Children* (Test des habiletés auditives et langagières chez l'enfant)

Hale Hancer
Erhan Kiziltan
Pinar Civak Tan
Derya Gokmen
Serhat Hayme
Suna Tokgoz Yilmaz

KEYWORDS
AUDITORY PROCESSING
SPEECH PROCESSING
AUDITORY DISCRIMINATION
SPEECH DISCRIMINATION
REACTION TIME

Hale Hancer[1], Erhan Kiziltan[2], Pinar Civak Tan[3], Derya Gokmen[4], Serhat Hayme[5], Suna Tokgoz Yilmaz[6]

[1]Department of Audiology, Baskent University, Ankara, TÜRKIYE

[2]Department of Physiology, Baskent University, Ankara, TÜRKIYE

[3]Department of Audiology and Speech Disorders, Ankara University, Ankara, TÜRKIYE

[4]Department of Biostatistics, Ankara University, Ankara, TÜRKIYE

[5]Department of Biostatistics and Medical Informatics, Erzincan Binali Yildirim University, Erzincan, TÜRKIYE

[6]Department of Audiology, Ankara University, Ankara, TÜRKIYE

Editor:
Victoria Duda

Abstract

Auditory processing and speech processing disorders negatively affect school-aged children. To minimize these negative effects, individuals in the risk group should benefit from the positive contribution of early intervention with a comprehensive evaluation. The aim of this study was to develop the Auditory and Speech Performance Test for Children and analyze its validity and reliability. In the development of the Auditory and Speech Performance Test for Children, discrimination and recognition subtasks were built for both auditory and speech performance. Meaningful and meaningless minimal pairs were used in the subtasks. A silent background was used for auditory performance, and noise stimuli were combined into minimal pairs for speech performance. Audiovisual materials were integrated into the finger-tapping test. The Auditory and Speech Performance Test for Children was administered to 307 children with typical development and to 80 children with specific learning disabilities. The Auditory and Speech Performance Test for Children calculated children's reaction times for pressing speed and accuracy of pressing the correct key. The data were analyzed for content, construct validity, internal consistency, and test-retest reliability. The content validity index value was found to be high (.89–1.0). The Auditory and Speech Performance Test for Children was explained as a two-factor model using exploratory factor analysis (eigenvalue = 1.92, total variance = 66.65%). It was found to be discriminative according to age, groups, subtests, and 27% bottom and top scores (all were significant at $p < .001$). Internal consistency (.77–.90) and test-retest values (.89–.93) of the Auditory and Speech Performance Test for Children in the total test scores were calculated within reliable values. In conclusion, we developed a valid and reliable screening tool for auditory and speech performance in children.

Abrégé

Les troubles de traitement auditif et de traitement des informations langagières ont des effets négatifs sur les enfants d'âge scolaire. Afin d'atténuer ces effets négatifs, il serait préférable que les enfants à risque bénéficient d'une évaluation complète des habiletés de traitement auditif et de traitement des informations langagières, ainsi que des avantages que procure l'intervention précoce. L'objectif de cette étude était de développer le *Auditory and Speech Performance Test for Children (*Test des habiletés auditives et langagières chez l'enfant) et d'en mesurer la validité et la fiabilité. Pendant la phase de développement de ce test, des sous-tâches de discrimination et de reconnaissance ont été conçues pour évaluer les habiletés auditives et langagières. Des paires minimales composées de mots et de non-mots étaient utilisées dans ces sous-tâches. L'évaluation des habiletés auditives a été effectuée à l'aide de paires minimales sans la présence de bruit de fond, tandis que les habiletés langagières ont été évaluées à l'aide de paires minimales présentées avec un bruit de fond. Du matériel audiovisuel a été intégré à un test de tapotement du doigt. Le *Auditory and Speech Performance Test for Children* a été utilisé auprès de 307 enfants au développement typique et 80 enfants ayant un trouble spécifique des apprentissages. Il a mesuré le temps de réaction avec lequel les enfants appuyaient sur une touche et la précision des touches enfoncées. Les données ont été analysées de façon à déterminer les validités de contenu et de construit, la cohérence interne et la fiabilité test-retest. Les résultats ont montré que la valeur de l'indice de validité du contenu était élevée (0,89–1,0). Les résultats de l'analyse factorielle exploratoire ont montré que le *Auditory and Speech Performance Test for Children* était expliqué par un modèle à deux facteurs (valeur propre = 1,92, variance totale = 66,65 %). Les résultats ont aussi montré que le *Auditory and Speech Performance Test for Children* permettait de discriminer selon l'âge, le groupe, la sous-tâche et 27 % des scores supérieurs et inférieurs (tous étaient significatifs à $p < 0,001$). Les valeurs calculées pour la cohérence interne (0,77–0,90) et la fidélité test-retest (0,89–0,93) se retrouvaient à l'intérieur de l'intervalle de fiabilité. En conclusion, nous avons développé un outil de dépistage valide et fiable pour évaluer les habiletés auditives et langagières des enfants

Auditory processing disorder (APD) refers to impairment in the perceptual processing of auditory information in the central auditory nervous system (American Speech-Language- Hearing Association [ASHA], 2005a), whereas *speech processing disorder* refers to difficulty interpreting or comprehending auditory information (Richard, 2017). Because these disorders are not well known, they are often confused with other neurodevelopmental disorders such as specific learning difficulties (SLD), attention deficits, and hyperactivity disorders (ASHA, 2005a, 2005b; Bellis, 2011; Chermak et al., 1997).

A comprehensive assessment of auditory and speech processing skills is important for intervention and follow-up (ASHA, 2005a; Bellis, 2011). APD test batteries are ignored because they are not functional owing to the need for professional experts and ineffective use of time (ASHA, 2005b). There is currently no gold standard for evaluating speech or auditory processing skills (Palana et al., 2022; Richard, 2017). Consequently, delayed diagnosis or misdiagnosis may occur due to the lack of clarity in the diagnostic protocol (Geffner & Ross-Swain, 2018). However, early intervention for processing skills is significant as it contributes positively to school achievement (Cacace & McFarland, 1998; DeBonis & Moncrieff, 2008), language and speech development (Barrazo et al., 2016; Moore, 2007), cognitive abilities (Tomlin et al., 2015), and psychosocial status (Kreisman et al., 2012).

Therefore, it is important to use screening tools to identify individuals at risk, such as those with hearing loss, otitis media, communication disorders, neurodevelopmental disorders, neurological disorders, hereditary predispositions, or premature birth, and direct them to a comprehensive evaluation for auditory and speech processing skills (Bellis, 2011; Geffner & Ross-Swain, 2018). When the scanning tools are examined, there are tools by which recognition and processing skills can be evaluated in many ways, if not as a whole (Goldman, 2015; Reynolds et al., 2005). These screening tools include performance-based measurements and observations of families and teachers (Geffner & Ross-Swain, 2018; Musiek & Chermak, 2007). Additionally, processing time and nature can be determined using electrophysiological tests (ASHA, 2005a). However, performing these procedures in clinical settings does not take place because of limited time, equipment availability, and patience of the individual. Therefore, behavioural tests are preferred in clinical settings. It has been argued that auditory processing can be evaluated using tonal and speech stimuli as behavioural measures (Chermak et al., 1997; Masters et al., 1998). It is believed that tonal tests do not reflect daily life auditory processing skills, so it is recommended that auditory processing skills be evaluated using speech stimuli (Katz, 2016). However, auditory processing cannot be measured exactly due to the effect of cognitive loads when a speech stimulus is used (Musiek & Chermak, 2007).

In the literature, there are behavioural test batteries that screen and provide a comprehensive evaluation of auditory and speech processing skills. These test batteries include both tonal and speech stimuli. For example, the Test for Auditory Processing Disorders (SCAN; Keith, 2000) is an evaluation tool for adults and children with APD. It consists of four subtests: Filtered Words, Auditory Figure Ground, Competing Words, and Competing Sentences. A standardized assessment tool called the Language Processing Test-3 (Richard & Hanner, 2005) is used to evaluate various language processing skills in people between the ages of 5 and 21 and includes six subtests: Associations, Categorization, Similarities, Differences, Multiple Meanings, and Attributes. The Screening Test for Auditory Processing (STAP; Yathiraj & Maggu, 2013) was designed to scan children for APD. It includes Speech Perception in Noise, Dichotic Consonant-Vowel, Gap Detection, and Auditory Memory subtests.

The differences in responses to auditory stimuli in noisy and quiet backgrounds can reveal reaction times for processing (Houben et al., 2013; Meister et al., 2018; Rönnberg et al., 2013). To evaluate processing skills, individuals can respond to stimuli by speaking, pointing out, using eye movements, or pressing a button in the presence or absence of background noise. (Geffner & Ross-Swain, 2018; Katz, 2016; Keith, 2000; Martin & Brownell, 2005; Musiek & Chermak, 2007). Verbal responses and reaction times can be recorded for auditory or speech processing skills (Holden et al., 2019; Meister et al., 2018). Meister et al. (2018) investigated the verbal reaction time during the conventional speech-in-noise test, grouping the participants according to age and listening status, then examining noise types and intelligibility levels, and they stated that verbal reaction time could be easily evaluated during conventional speech audiometry.

In the literature, it is reported that the prevalence of APD in children ranges from 73% to 96% (Wilson & Arnott, 2013), with a male-to-female ratio of 2:1 (Chermak et al., 1997). In adults aged 55 and above, the prevalence varies from 23% to 76% (Golding et al., 2004). Considering the profound impact of auditory and speech processing skills on an individual's quality of life, a comprehensive assessment and early intervention are important when there is suspicion of a disorder in these skills (Geffner & Ross-Swain, 2018; Musiek & Chermak, 2007).

Therefore, there is a pressing need to develop an assessment tool that can rapidly screen individuals and is easily accessible to clinicians. This assessment tool would not only contribute to the positive outcomes of early intervention but also support comprehensive therapy programs. Furthermore, focusing on processing skills in more detail, especially in individuals with cognitive and sensory impairments, central auditory processing, and language and speech disorders, could contribute to the functional use of these dimensions, which received limited or no attention in therapy programs (Cacace & McFarland, 1998; Geffner & Ross-Swain, 2018; Katz, 2016)

A review of the literature reveals that screening tests, both observation-based and performance-based, are available for assessing auditory and speech processing. However, performance-based screening tests may have limitations in evaluating everyday life conditions and may require clinical settings for assessment. Consequently, there is a need for a screening test that evaluates processing based on real-life challenges. Furthermore, there is no experimental behavioural test to measure processing speed, even at the screening level in the literature. Because there are controversies in the evaluation of processing skills, it is important to develop performance screening tools that will indirectly reflect processing skills (Smoski et al., 1998). Considering auditory discrimination and recognition skills as performance tests (Archbold et al., 1998; Smoski et al., 1998), this study aimed to develop a screening method to determine auditory and speech skills.

As there is no gold standard for diagnosing APD, this study included individuals with various disorders that are thought to have similar findings as APD, instead of only those diagnosed with APD. Therefore, in addition to participants with typical development, individuals diagnosed with SLD, whose auditory processing abilities are known to be affected (Dawes & Bishop, 2010; Hämäläinen et al., 2013; King et al., 2003), also participated in this study. The literature indicates that individuals with typical development and SLD are commonly included as participants in APD studies. For instance, the development of the Comprehensive Test of Phonological Processing-2 (CTOPP-2; Wagner et al., 1999) involved 1900 participants aged 4 to 24 years with typical development, while the Auditory Skills Assessment (Geffner & Goldman, 2010) was conducted with 475 typical participants for test development. Domitz and Schow (2000) included individuals diagnosed with SLD as an atypical group during the development of the Multiple Auditory Processing Assessment-2 (MAPA-2), and Martin et al. (2018) included individuals diagnosed with SLD during the development of the Test of Auditory Processing-4.

The primary aim of this study was to develop a screening test called the Auditory and Speech Performance Test for Children (ASPT-C), which evaluates the auditory and speech performance of children 7.0 to 10;12 years old, with the parameters of recording correct/false and millisecond-rate reactions to meaningful and meaningless rhyming word pairs in the presence or absence of background noise. When psychoeducational tests developed to assess processing skills are examined, various validity parameters are evaluated in the validity study of these tests. For example, Webster (2009) assessed the content and construct validity of the Test of Information Processing Skills in a validity study, and Martin and Brownell (2005) examined the correlation between participants' test scores and their age and intelligence scores in a validity study of the Test of Auditory Processing-3. In this study, we conducted an assessment of the content, construct validity, and reliability of the ASPT-C.

Method

Participants

School-age children with SLD who complained of difficulty understanding in noise and children with typical development participated in this study. School-age children in the schools of the districts of Ankara, whose parents' consent was obtained and who participated voluntarily, were included in the study.

A total of 387 children participated in this methodological study: 307 children with typical development (Group 1) and 80 children with SLD (Group 2). The reason for the onset age of 7.0 was literacy and auditory processing maturity (Jerger & Musiek, 2000; Moore et al., 2011). This study was limited to elementary school children due to the challenges encountered in accessing schools as a result of the COVID-19 pandemic. Participants were selected using stratified sampling according to age, gender, and district.

Because the nonstandardized diagnostic tools of APD and the literature indicate that auditory processing is also affected in SLD (Alles et al., 2011; ASHA, 2005a; Bellis, 2011; Eggermont, 2015; Moore et al., 2011; Sharma et al., 2009), 80 school-age children with SLD who had complaints of difficulty understanding in noisy environments were included in the study by stratified sampling according to age and gender (Group 2). To determine the complaint of difficulty understanding noise, a data collection form prepared by the authors was used due to the lack of a valid and reliable scale for these ages in Türkiye. The data form included questions related to comprehension of speech, understanding of speech in noisy environments, following instructions, and history of language delay. The data form was reviewed by five clinical audiologists working in research hospitals (3 females, age *M* = 27.3, *SD* = 2.8, professional experience *M* = 6.2 years, *SD* = 1.3) to ensure content validity.

The inclusion criteria for Group 1 were voluntary participants who were right-dominant handed, stated to have no disorders by their teachers or family, literate, and had consent from their families; Group 2 included voluntary participants who were right-dominant handed, literate, and had no diagnosis except for a diagnosis of SLD, complained of understanding in noise, and had consent from their families.

To assess the reading abilities of the groups, reading passages were read aloud to the children for 1 min, according to age-specific texts prepared by the authors and linguists. Participants who were at the age cutoff point were included in the study (for further details, see Erden et al., 2002).

Development of ASPT-C Items

The standards for reporting diagnostic accuracy (Bossuyt et al., 2015) guideline, in addition to reference sources (Boateng et al., 2018; de Vet et al., 2011; Streiner et al., 2015), were reviewed to guide both the study design and the development of the ASPT-C items.

After the literature review, the following four tasks were defined and developed for the ASPT-C. While developing the tasks of the ASPT-C, certain steps of listening skills were considered to reflect the recognition of speech in daily life (American Academy of Audiology [AAA], 2010; Archbold et al., 1998). These steps include detection, discrimination, recognition, and comprehension (Erber, 1975). The detection level was determined by the presence or absence of sounds. The discrimination level focused on similarities and differences in the sound (AAA, 2010; Estabrooks et al., 2016). Accordingly, it can be measured by showing or repeating what was heard between two similar sounds or by determining whether the minimal pairs were the same (Meinzen-Derr et al., 2007). The AAA (2010) recommended assessing the discrimination level of listening in the evaluation of auditory processing skills. In this context, the ability to distinguish was also examined using the ASPT-C. The recognition level is defined as the ability to perceive a stimulus under difficult listening conditions (Erber, 1975; Meinzen-Derr et al., 2007). Based on this definition, in the ASPT-C, meaningless items were created to simulate difficult listening conditions and to control for executive functions, working memory, and the listening effort effect in the evaluation of speech processing skills (Danneels et al., 2021; DeBonis, 2015).

In the evaluation of processing skills, it is recommended to evaluate the discrimination of speech in noise to reflect daily life (Archbold et al., 1998; Bellis, 2011; Yathiraj & Vanaja, 2018). When the quality of the signal decreases due to factors such as noise or hearing loss, speech processing can become more challenging, particularly when the language used is complex or the message content is less familiar (Akeroyd, 2008; Danneels et al., 2021; Libben et al., 2020). Therefore, noise stimuli were added to the speech-processing tasks in the ASPT-C, with noisy backgrounds that are often encountered in daily life, while presentation in quiet backgrounds was used for auditory performance assessment.

For the auditory discrimination subdimension, measures included the reaction times and the number of correct answers to meaningful monosyllabic rhyming words in silence; for the auditory recognition subdimension, the reaction times and the number of correct answers to meaningless monosyllabic rhyming words in silence; for the speech discrimination subdimension, the reaction times and the number of correct answers to meaningful monosyllabic rhyming words in background noise; and for the speech recognition subdimension, the reaction times and the number of correct answers to meaningless monosyllabic rhyming words in background noise.

Task items were prepared according to the manner, place, and voicing of phonemes and, as monosyllabic minimal pair words, changed sounds at the beginning and end of the word. When preparing meaningless monosyllabic word pairs, vowel-distributed words were chosen to prevent intelligibility skills and inferences from affecting the data (Steadman & Sumner, 2018). Prepared words were collected from a pool and checked by a linguist. Each task consisted of 15 items, and 60 items were developed. (see **Table 1**).

Audiovisual Materials

A male voice was used to record one of the words in each word pair in a sound studio. In this study, a male voice was used as the auditory stimulus following a comprehensive review of reference tests (Katz, 2016; Keith, 2000; Reynolds et al., 2005). The stimulus was presented at 50 dB HL with background noise added for speech tasks, and the signal-to-noise ratio was set to 0. This selection of 50 dB HL as the listening level was made because the most comfortable listening level would be at least 15 dB above background noise (Kobayashi et al., 2007; Ueda & Tanaka, 2020). In the ASPT-C, each word pair was presented twice, with and without background noise, and the answers were randomized. For instance, in

Table 1

Auditory and Speech Performance Test for Children Task Items

Number	Discrimination performance		Recognition performance	
	Auditory discrimination	Speech discrimination	Auditory recognition	Speech recognition
1	**bil**-pil	bil-**pil**	fim- **fom**	**fim**- fom
2	**var**-far	**var**-far	min-**mun**	min-**mun**
3	cam-**çam**	**cam**-çam	liş-**laş**	liş-**laş**
4	**tüh**-tüy	**tüh**-tüy	**yeç**-yıç	**yeç**-yıç
5	kıs-**kız**	kıs-**kız**	**sem**-sam	sem-**sam**
6	buz-**muz**	**buz**-muz	baf-**bif**	**baf**-bif
7	taç- **saç**	taç- **saç**	**çum**-çem	çum-**çem**
8	**et**-ek	et-**ek**	**pez**-pız	**pez**-pız
9	**bar**- bal	bar- **bal**	nef-**naf**	**nef**-naf
10	harf-**harp**	**harf**-harp	**fap**-fip	fap-**fip**
11	**çiz**-diz	çiz-**diz**	kuç-**keç**	**kuç**-keç
12	**pas**-tas	**pas**-tas	dız-**döz**	**dız**-döz
13	del-**gel**	del-**gel**	**ron**-rün	**ron**-rün
14	**hoş**- loş	**hoş**- loş	kım-**kem**	kım-**kem**
15	**puf**-pus	**puf**-pus	**şit**-şat	şit-**şat**

Note. Boldface values represent the correct answers.

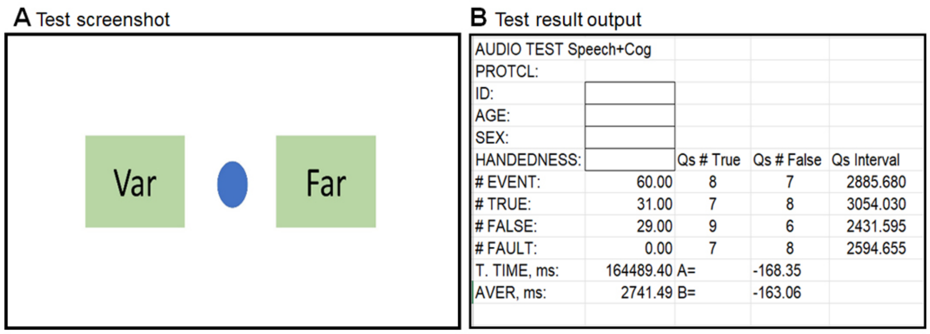
the word pair "fan-van," the correct answer was "fan" in the speech recognition task, while in the speech discrimination task, it could be "fan" or "van."

The type of background noise used in the study was "restaurant ambient." The reason for the use of this noise type was the reliability of the measurement by considering selective attention in the speech stimulus, which is sufficient to measure central auditory processing and is more convenient in daily life (Brungart & Simpson, 2007; Evans et al., 2016; Holube, 2011). The background noise (Stephan, n.d.) was obtained from the soundbible.com website where free audio files are available. The necessary arrangements were made as mentioned in the literature (Leccese et al., 2015; Zokoll et al., 2015) before superimposing the noise on the male voice.

Sound recordings were performed in a sound studio with the help of a sound technician. The devices used for recording and editing were Sound Recorder (Shure PG42-LC), a condenser microphone and speaker (Sony), and Sound Forge PRO 10 (Sony). The soundboard was a Soundcraft FX16. During sound recording, the microphone distance from the speaker was maintained at a minimum of 30 cm by positioning the microphone for optimal directionality; mono recordings were digitized at a sampling rate of 44.1 kHz with 16-bit amplitude resolution. The Audition 11.0 (Adobe) program was used to combine the stimuli prepared in the studio with the noise stimulus. The program used in arranging the noise characteristics and preventing sound explosions was based on Leccese et al.'s (2015) study, which utilized a pop-up filter.

The test items were prepared using written visuals. In the test, hand-eye coordination was not considered, and the word pairs on the screen were positioned perceptually on the right and left simultaneously in the middle of the screen so that the visual sense did not affect the results of the test (Jain et al., 2015). Participants were asked to perform this procedure by looking at the blue dot between two written words (Amini Vishteh et al., 2019). Participants could change the displayed word pairs by pressing a predefined button (see **Figure 1**).

Figure 1



AUDIO TEST Speech+Cog				
PROTCL:				
ID:				
AGE:				
SEX:				
HANDEDNESS:		Qs # True	Qs # False	Qs Interval
# EVENT:	60.00	8	7	2885.680
# TRUE:	31.00	7	8	3054.030
# FALSE:	29.00	9	6	2431.595
# FAULT:	0.00	7	8	2594.655
T. TIME, ms:	164489.40	A=	-168.35	
AVER, ms:	2741.49	B=	-163.06	

Screenshots of Auditory and Speech Performance Test for Children (ASPT-C) and data output

Note. Panel A: Screenshot of ASPT-C. Minimal pairs were located on the right and left sides of the midline of the screen. Participants listened to the stimulus while looking at the blue dot and pressed the keyboard key assigned to the appropriate written stimulus for the auditory stimulus. Panel B: Data output. The output included scores for data (reaction times and accuracy on subtasks) and participant characteristics.

Software

Auditory and visual stimuli were embedded into the digital finger-tapping test battery (Kiziltan et al., 2006), which measures tapping performance with a high time resolution and saves the data on the hard drive of a computer for analysis. The software was developed for use on a standard personal computer without the need for auxiliary hardware. The system used a read-time stamp counter, which is a powerful benchmark introduced by Intel in Pentium processors (Intel, Santa Clara, CA). Therefore, in an IBM-compatible personal computer with at least 1 GHz CPU, it is possible to reach a time resolution of milliseconds in measuring finger-tapping tasks (Aydin et al., 2016). With the new version of the software, it became possible to measure reaction times and responses to audiovisual stimuli. Participants were asked to press the predefined key on the computer keyboard as soon as they answered the audiovisual task. The test module automatically calculated the average reaction times and correct answers for the tasks and total test (see **Figure 1**).

Procedure

This study was conducted with ethical approval from the Ankara University Directorate (17.12.2021, I11-693-21). The study followed the principles of the Declaration of Helsinki with specific measures implemented to ensure participant confidentiality, informed consent, and fair treatment. Data collection was authorized by the Ankara Governorship Directorate of National Education (20.11.2020, 14588481-605.99-E.17020975).

All tests were conducted on a Lenovo Desktop-UTOJTA5 notebook with an Intel Core i5-4210U CPU @ 1.70 GHz 2.40 GHz processor and 64-bit operating system. The tests were administered in quiet locations within schools, such as the library or manager's office.

Before the main procedure, a preliminary study was conducted with 20 children with typical development (10 females; age M = 8.52, SD = 2.1) to ensure that the instructions, audiovisual materials, and test items were appropriate.

During the main procedure, each participant sat comfortably in front of the computer screen and was instructed to press a predefined keyboard key in response to each task. This instruction was presented to each participant.

Statistical Analyses

Regarding the psychometric properties of the test, answers to the research questions about the validity and reliability of the test were sought.

In the validity study, the content validity ratio (CVR) and content validity index (CVI) were determined for content validity. A correlation matrix suitable for exploratory factor analysis (EFA), Kaiser-Meyer-Olkin (KMO), Bartlett's sphericity test, and common variance values were determined and EFA was performed. As the data did not have a normal distribution for discriminant validity, differences between groups were calculated using the Mann-Whitney U test and multigroup differences were calculated using the Kruskal-Wallis test. As there was a difference, the Mann-Whitney U test (with the corrected Bonferroni test) was used to determine the groups from which the difference originated. Student's *t* test was used for item discrimination.

For internal consistency, Cronbach's alpha and the Kuder Richardson-20 (KR-20) reliability coefficient were used, and for test-retest reliability, the intraclass correlation coefficient with the two-way mixed model was analyzed.

Results

Participants

Of the participants, 39.8% (*n* = 154) were female and 60.2% (*n* = 233) were male. When analyzed by age, 17.05% (*n* = 66) were between 7.0 and 7;12 years old, 30.24% (*n* = 117) were between 8.0 and 8;12 years old, 29.20% (*n* = 113) were between 9.0 and 9;12 years old, and 23.51% (*n* = 91) were between 10.0 and 10;12 years old. According to socioeconomic status, 33.3% (*n* = 129) were classified as low, 34.7% (*n* = 134) as moderate, and 32% (*n* = 124) as high. According to grade level, 16.80% (*n* = 65) were in the 2nd grade, 29.71% (*n* = 115) in the 3rd grade, 29.97% (*n* = 116) in the 4th grade, and 23.52% (*n* = 91) in the 5th grade (see **Table 2**).

Content Validity

Expert opinion was sought to determine the content validity of the tests recorded by the CVR and CVI. Accordingly, the CVR for the test items were calculated with the formulation [G/(N/2)]–1 (where G = number of experts scoring, 4+5/3+4+5 and *N* = total number of experts). CVI is the average CVR value of the items remaining in the pool (Lawshe, 1975). For expert opinions, the CVR of test items were 1.0 for the Auditory Discrimination task; .94 for the Speech Discrimination task; .89 for the Auditory Recognition task; .92 for the Speech Recognition task. CVI values were found 1.0 in the Auditory Discrimination task; .94 in the Speech Discrimination task; .89 in the Auditory Recognition task; and .92 in the Speech Recognition task.

Extraction of Factors

EFA was employed to assess the factors that fit the items. (Boateng et al., 2018; DeVellis & Thorpe, 2021). Before applying EFA, the suitability of the items for the factor structures of the test was checked using a correlation matrix. According to the test tasks, the correlation matrix of the test items had a factor load greater than .30.

Table 2

Participant Characteristics

Characteristics	Group 1		Group 2		Total	
	n	%	*n*	%	*n*	%
Gender						
Female	123	40.00	31	38.75	154	39.80
Male	184	60.00	49	61.25	233	60.20
Ages in years; months						
7;0–7;12	53	17.26	13	16.25	66	17.05
8;0–8;12	92	29.96	25	31.25	117	30.24
9;0–9;12	90	29.32	23	28.75	113	29.20
10;0–10;12	72	23.46	19	23.75	91	23.51
Socioeconomic status						
Low	103	33.60	26	32.50	129	33.30
Middle	105	34.20	29	36.25	134	34.70
High	99	32.20	25	31.25	124	32.00
Grade						
2	52	16.94	13	16.25	65	16.80
3	90	29.32	25	31.25	115	29.71
4	95	30.94	21	26.25	116	29.97
5	70	22.80	21	26.25	91	23.52

The suitability of the ASPT-C data for factor analysis was examined using the KMO coefficient and Bartlett's sphericity test. In this study, the KMO value was .79, and Bartlett's Sphericity test yielded significant results ($x^2(6)$ = 717.594, *p* < .001). After determining whether the test was suitable for factor analysis, the common variances of the test items were examined and the values were found to be sufficient for the factor analysis. ASPT-C gathered two factors with an eigenvalue of 1.92, accounting for 66.65% of the total variance. The first factor was named Accuracy and the second factor was named Reaction Time (see **Table 3**).

Item Discrimination

To evaluate the item discrimination of the ASPT-C, the difference between the scores in the top 27% and bottom 27% was examined. For the Accuracy subtest, a significant difference was found between the bottom 27% (*n* = 104, *M* = 50.24, *SD* = 5.91) and top 27% (*n* = 104, *M* = 58.44, *SD* = 0.57), *t*(207) = −14.13, *p* < .001, *r* = 1.95. When the Reaction Time data were examined, a significant difference was found between the bottom 27% (*n* = 104, *M* = 1892.33, *SD* = 110.69) and the top 27% (*n* = 104, *M* = 2667.29, *SD* = 258.91), *t*(207) = −28.09, *p* < .001, *r* = 3.89; see **Table 3** for more detail).

Construct Validity

Construct validity was assessed through differentiation based on "known groups" by analyzing the differences among groups, ages, and task scores.

To compare the Accuracy and Reaction Time scores between Group 1 (*M* = 56, *SD* = 3.8 for Accuracy, *M* = 2175.1, *SD* = 285.8 for Reaction Time) and Group 2 (*M* = 51.4, *SD* = 4.9 for Accuracy, *M* = 2483, *SD* = 375.5 for Reaction Time), both the total test and tasks were considered (see **Table 4**). The results showed that the Accuracy and Reaction Time scores were statistically significant (*U* = 3894.5, *z* = −9.48, *p* < .001, *r* = .48 for Accuracy total score; *U* = 6339, *z* = −6.66, *p* < .001, *r* = .33 for Reaction Time total score).

When the background noise in the discrimination subtasks of Group 1 was examined, the difference between the presence of background noise (*M* = 13.5, *SD* = 1.2) and

Table 3

Total Variance Results and Eigenvalue Coefficient of Auditory and Speech Performance Test for Children

Tasks	Factor loading		Top 27% and bottom 27% group comparison
	1	2	*t*
Factor 1: Accuracy, Eigenvalue = 4.74, Total variance = 37.8%			
Auditory Discrimination	−.148	**.686**	−16.25* (df=208)
Speech Discrimination	−.236	**.611**	−21.23* (df=208)
Auditory Recognition	−.068	**.782**	−15.37* (df=208)
Speech Recognition	−.153	**.689**	−27.77* (df=208)
Factor 1's total score	**−.184**	**.887**	**−14.13* (df=207)**
Factor 2: Reaction Time, Eigenvalue = 1.92, Total variance = 28.85%			
Auditory Discrimination	**.801**	−.122	−24.22* (df=207)
Speech Discrimination	**.809**	−.210	−24.67* (df=207)
Auditory Recognition	**.820**	−.163	−19.40* (df=207)
Speech Recognition	**.849**	−.198	−23.27* (df=207)
Factor 2's total score	**.975**	**−.208**	**−28.09* (df=207)**
Total Eigenvalue = 1.92, Total variance = 66.65%			

Note. Factor analysis after varimax rotation. Boldface values represent primary loading associated with each factor.
**p* < .001

Table 4

Comparison Between Groups According to Accuracy and Reaction Time

Task	Subtask	Group 1		Group 2		*U*	*z*	*r*
		M ± *SD*	*z*-scores	*M* ± *SD*	*z*-scores			
Accuracy	Auditory Disc.	14.4 ± 0.9	−9.81*	13.3 ± 1.9	−5.48*	7273.5	−6.21*	.31
	Speech Disc.	13.5 ± 1.2		11.7 ± 1.7		4574.5	−8.86*	.53
	Auditory Recog.	14.6 ± 0.7	−10.58*	13.8 ± 1.5	−5.86*	8332.5	−5.22*	.26
	Speech Recog.	13.8 ± 1.0		12.7 ± 1.8		7355.5	−5.74*	.29
	Total score	56 ± 3.8		51.4 ± 4.9		3894.5	−9.48*	.48
Reaction time (ms)	Auditory Disc.	2169.3 ± 330.6	−10.65*	2393.9 ± 478.9	−6.20*	8715	−4.00*	.20
	Speech Disc.	2379.3 ± 385.8		2759.9 ± 472.4		6325	−6.68*	.33
	Auditory Recog.	1935.2 ± 316.5	−13.23*	2277.1 ± 488.5	−5.25*	6910	−6.02*	.30
	Speech Recog.	2206.2 ± 330.7		2500.9 ± 441.8		7271	−5.62*	.28
	Total score	2175.1 ± 285.8		2483 ± 375.5		6339	−6.66*	.33

Note. Disc = Discrimination; Recog = Recognition. z-scores are based on the presence of noise stimulus in subtests.
* *p* < .001

absence of background noise ($M = 14.4$, $SD = 0.9$) in the Accuracy subtest, and the presence of background noise ($M = 2379.3$, $SD = 385.8$) and the absence of background noise ($M = 2169.3$, $SD = 330.6$) in the Reaction Time subtests were significant ($z = -9.81$, $p < .001$, $r = .56$ for Accuracy; $z = -10.65$, $p < .001$, $r = .6$ for Reaction Time). When considered as recognition subtasks, the scores between the presence of background noise ($M = 13.8$, $SD = 1.0$) and absence of background noise ($M = 14.6$, $SD = 0.7$) in the Accuracy subtest and between the presence of background noise ($M = 2206.2$, $SD = 330.7$), and the absence of background noise ($M = 1935.2$, $SD = 316.5$) in the Reaction Time subtest was also significant ($z = -10.58$, $p < .001$, $r = .6$ for Accuracy; $z = -13.23$, $p < .001$, $r = .75$ for Reaction Time). When the background noise in the discrimination subtasks of group 2 was examined, the difference between the presence of background noise ($M = 11.7$, $SD = 1.7$) and absence of background noise ($M = 13.3$, $SD = 1.9$) in the Accuracy subtest, and presence of background noise ($M = 2759.9$, $SD = 472.4$) and absence of background noise ($M = 2393.9$, $SD = 478.9$) in the Reaction Time subtest were significant ($z = -5.48$, $p < .001$, $r = .6$ for Accuracy; $z = -6.20$, $p < .001$, $r = .69$ for Reaction Time). Likewise, in the recognition subtest, the scores between the presence of background noise ($M = 12.7$, $SD = 1.8$) and absence of background noise ($M = 13.8$, $SD = 1.5$) in the Accuracy subtest, and between the presence of background noise ($M = 2500.9$, $SD = 441.8$) and the absence of background noise ($M = 2277$, $SD = 488.4$) in Reaction Time were found to be significant ($z = -5.86$, $p < .001$, $r = .65$ for Accuracy; $z = -5.25$, $p < .001$, $r = .58$ for Reaction Time; see **Table 4**). Accordingly, when the noise stimulus was added, the accuracy decreased and the reaction time was delayed.

When analyzing the ASPT-C scores according to age, significant differences were found in the Accuracy scores of Group 1 ($H(3) = 49.20$, $p < .001$). Pairwise comparisons showed differences between the ages of 7 and 9 ($z = -52.18$, $p < .001$, $r = 4.3$), between the ages of 7 and 10 ($z = -88.09$, $p < .001$, $r = 8$), between the ages of 8 and 9 ($z = -44.26$, $p < .001$), and between the ages of 8 and 10 ($z = -80.16$, $p < .001$, $r = 6.2$). Significant differences were also found in the Reaction Time scores between ages ($H(3) = 110.64$, $p < .001$). Pairwise comparisons showed differences between the ages of 7 and 9 ($z = 68.19$, $p < .001$, $r = 5.6$), between the ages of 7 and 10 ($z = 114.28$, $p < .001$), between the ages of 8 and 9 ($z = 40.75$, $p < .001$, $r = 3.7$), between the ages of 8 and 10 ($z = 116.84$, $p < .001$, $r = 9$), and between the ages of 9 and 10 years ($z = 76.09$, $p < .001$, $r = 6$). In Group 2, there were no significant differences in Accuracy and Reaction Time scores between ages ($H(3) = 8.287$, $p = .040$ for Accuracy; $H(3) = 4.931$, $p = .177$ for Reaction Time).

In summary, ASPT-C has content validity. It has a two-factor structure: Reaction Time and Accuracy. The ASPT-C is distinctive according to groups, age, and subtest.

Reliability

Cronbach's alpha and KR-20 were used to assess the internal consistency and reliability of the ASPT-C test. The Cronbach's alpha values of the ASPT-C's Reaction Time tasks ranged from .77 to .90, indicating moderate to high internal consistency (see **Table 5**). The total scale had a

Table 5

Auditory and Speech Performance Test for Children Cronbach's Alpha (α) Values for Internal Consistency

Tasks	*M* ± *SD*	α (*n* = 387)	ICC (*n* = 120)	Confidence interval
Auditory Discrimination	2215.69 ± 376.64	.77	.67	.47–.79
Speech Discrimination	2458.03 ± 433.03	.78	.76	.62–.84
Auditory Recognition	2005.86 ± 384.03	.90	.80	.68–.87
Auditory Recognition	2267.14 ± 375.37	.80	.74	.58–.84
Total score	2238.75 ± 330.47	.93	.93	.90–.96

Note. ICC = intraclass correlation coefficient

Cronbach's alpha of .93, indicating high reliability. Similarly, the KR-20 coefficient for the Accuracy tasks was .70, suggesting moderate reliability (see **Table 6**). However, the KR-20 values for the subscores ranged from .20 to .56, indicating low internal consistency.

Test-retest reliability was assessed using the intraclass correlation coefficient, which ranged from .67 to .80 for Reaction Time and .49 to .76 for Accuracy. The correlation coefficients were .93 for Reaction Time ($F(79) = 11.36$, $p < .001$) and .89 for Accuracy ($F(79) = 16.06$, $p < .001$), indicating high reliability for the total scores.

In summary, the subtasks of Accuracy and Reaction Time were moderately reliable, and the total scores of Accuracy and Reaction Time were highly reliable.

Discussion

In this study, 307 school-age children with typical development and 80 school-age children with SLD were evaluated for auditory and speech performance using the ASPT-C. The test measured correct answer scores and reaction times in milliseconds in response to meaningful and meaningless minimal pairs, in both the presence and absence of background noise. Content validity, factor extraction, item discrimination, construct validity, internal consistency, and test-retest reliability of the ASPT-C were analyzed.

When developing a scale or test, it is important to determine validity and reliability. EFA, one of the methods used to identify factor loads before establishing the construct validity of the assessment tool, was conducted (DeVellis & Thorpe, 2021). Upon reviewing the literature, EFA was used for the construct validity of the STAP (Yathiraj & Maggu, 2013), which had three factors, and Domitz and Schow (2000) revealed that MAPA had four factors. In this study, we developed a test to explain these two factors by performing EFA. The ASPT-C consisted of four subtasks, each with two response categories (accuracy and reaction times). These response categories accounted for two-factor loading, and demonstrated the construct validity of the ASPT-C.

To enhance the construct validity of the ASPT-C, participant scores in the lower and upper 27% groups were compared, revealing item discrimination. Item discrimination is used to strengthen the construct validity of the items in scales and/or tests (Johnston et al., 2014). When reviewing the literature related to the assessment of auditory or speech processing skills, no scale/test was found that demonstrated item discrimination. In this regard, we introduced a test with established item discrimination values in the literature.

A comparison of the test with an atypical group is also an important parameter for construct validity. Schow et al. (2021) included participants diagnosed with speech difficulties, attention-deficit/hyperactivity disorder, and dyslexia in the MAPA-2 test, as well as individuals with normal development in their study. Keith (2000) included individuals diagnosed with APD as both a typical developmental and atypical group in a validity test for children with APD (SCAN 3-C). In this study, we included children with SLD who had complaints of difficulty understanding in noisy environments. By comparing the two groups, statistically significant differences were found in test scores between children with typical development and those with SLD. This result demonstrated that ASPT-C reveals the difference between the groups.

To demonstrate the construct validity, age groups, which are thought to affect auditory and speech processing skills, were compared. When these findings were examined, differences were observed among the age groups. It could be seen that as age increased, both the reaction speed and the number of correct responses increased. This finding

Table 6

Auditory and Speech Performance Test for Children KR-20 Values for Internal Consistency

Tasks	*M* ± *SD*	KR-20 coefficient (*n* = 387)	ICC (*n* = 120)	Confidence interval
Auditory Discrimination	14.18 ± 1.26	.56	.56	.32–.72
Speech Discrimination	13.12 ± 1.53	.39	.76	.63–.84
Auditory Recognition	14.43 ± 0.98	.44	.68	.50–.79
Auditory Recognition	13.53 ± 1.27	.20	.49	.20–.68
Total score	55.11 ± 4.44	.70	.89	.79–.94

Note. ICC = intraclass correlation coefficient

confirms the hypothesis that processing skills improve with age (Jerger & Musiek, 2000; Moore et al., 2011).

Children with auditory and speech processing disorders may encounter difficulties in understanding instructions in noisy environments in their daily lives. Therefore, noise stimuli are frequently employed to assess APD. In this study, we structured the ASPT-C subtasks based on the presence or absence of noisy stimuli. We found an increase in reaction times and incorrect responses with background noise, both among children with typical development and those with SLD. When comparing the groups, we found that children with SLD struggled more with these tasks. Ferenczy et al. (2022), Koiek et al. (2018), and Warrier et al. (2004) found significant differences in comprehension scores of individuals with SLD under noisy conditions. The results of this study demonstrate the effectiveness of noise stimuli in measuring processing skills and contribute to the extension of findings in the literature.

Internal consistency and intraclass correlation coefficients were calculated to examine the reliability of the ASPT-C. In the literature, reliability analyses of APD tests show highly reliable results for certain tests and tasks. However, some of these tests may have low reliabilities. In the reliability study of the Feather Squadron Test (Barker & Purdy, 2016), scores varied between 17.3 and 90.8 depending on the tasks. The reliability of CTOPP-2 (Wagner et al., 1999) tasks was found to have test-retest correlations between .75 and .92 for core tasks, between .76 and .86 for composites, and .73 and .75 for tasks. These low or medium reliability results were attributed to factors such as a lack of maturation in auditory processing, difficulty in directing attention, and listening effort. The reliability of the ASPT-C was also examined using internal consistency and stability parameters and was found to be moderately or highly reliable. This result is likely to have affected the reliability of the 7-year-old age group. In addition, we believe that the test was performed on a computer and that the attention of the participants, listening effort, and age were important factors.

The ASPT-C has demonstrated its potential for assessing auditory and speech processing skills in children. Consequently, it is believed that this test can be employed to screen for auditory and speech processing disorders and facilitate early intervention. ASPT-C offers practicality and ease of use as a noteworthy advantage. Additionally, this study is anticipated to contribute to further research in the field of auditory and speech processing skills. Although our findings provide insights into the impact of auditory processing skills on individuals with SLD, further research is necessary to generalize these results.

The limitations of this study are that the study sample was limited to Türkiye, and the ASPT-C was performed with a single computer to ensure proper calibration of the test.

Finally, we conclude that the ASPT-C is a valid and reliable screening test that can be applied to children aged 7.0–10.12 years. We recommend performing the test with a larger sample size, comparing the results with those of different groups of people with varying needs and ages, and using a device specifically designed and calibrated for the test to avoid calibration problems. Additionally, it is recommended to compare the scores obtained from different processing tests with those obtained from the ASPT-C.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47*(sup2), S53–S71. https://doi.org/10.1080/14992020802301142

Alles, R., Bamiou, D., Batchelor, L., Campbell, N. G., Canning, D., Grant, P., Luxon, L., Moore, D., Murray, P., & Nairn, S. (2011, March 31). *British Society of Audiology position statement: Auditory processing disorder (APD)*. https://eprints.soton.ac.uk/338076/1/BSA_APD_PositionPaper_31March11_FINAL.pdf

American Academy of Audiology. (2010, August). *Diagnosis, treatment and management of children and adults with central auditory processing disorder* [Clinical practice guidelines]. https://audiology-web.s3.amazonaws.com/migrated/CAPD%20Guidelines%208-2010.pdf_539952af956c79.73897613.pdf

American Speech-Language-Hearing Association. (2005a). *Central auditory processing disorder*. https://www.asha.org/Practice-Portal/Clinical-Topics/Central-Auditory-Processing-Disorder/

American Speech-Language-Hearing Association. (2005b). *(Central) auditory processing disorders—The role of the audiologist* [Position statement]. https://doi.org/10.1044/policy.PS2005-00114

Amini Vishteh, R., Mirzajani, A., Jafarzadehpour, E., & Darvishpour, S. (2019). Evaluation of simple visual reaction time of different colored light stimuli in visually normal students. *Clinical Optometry, 11*, 167–171. https://doi.org/10.2147%2FOPTO.S236328

Archbold, S., Lutman, M. E., & Nikolopoulos, T. (1998). Categories of auditory performance: Inter-user reliability. *British Journal of Audiology, 32*(1), 7–12. https://doi.org/10.3109/03005364000000045

Aydin, L., Kiziltan, E., & Gundogan, N. U. (2016). Polyphasic temporal behavior of finger-tapping performance: A measure of motor skills and fatigue. *Journal of Motor Behavior, 48*(1), 72–78. https://doi.org/10.1080/00222895.2015.1040111

Barker, M. D., & Purdy, S. C. (2016). An initial investigation into the validity of a computer-based auditory processing assessment (Feather squadron). *International Journal of Audiology, 55*(3), 173–183. https://doi.org/10.3109/14992027.2015.1074734

Barrozo, T. F., Pagan-Neves, L. D. O., Vilela, N., Carvallo, R. M. M., & Wertzner, H. F. (2016). The influence of (central) auditory processing disorder in speech sound disorders. *Brazilian Journal of Otorhinolaryngology, 82*(1), 56–64. https://doi.org/10.1016/j.bjorl.2015.01.008

Bellis, T. J. (2011). *Assessment and management of central auditory processing disorders in the educational setting: From science to practice*. Plural Publishing.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J.G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ, 351*, Article 5527. https://doi.org/10.1136/bmj.h5527

Brungart, D. S., & Simpson, B. D. (2007). Cocktail party listening in a dynamic multitalker environment. *Perception & Psychophysics, 69*(1), 79–91. https://doi.org/10.3758/BF03194455

Cacace, A. T., & McFarland, D. J. (1998). Central auditory processing disorder in school-aged children: A critical review. *Journal of Speech, Language, and Hearing Research, 41*(2), 355–373. https://doi.org/10.1044/jslhr.4102.355

Chermak, G. D., Musiek, F. E., & Craig, C. H. (1997). *Central auditory processing disorders: New perspectives.* Singular.

Danneels, M., Degeest, S., Dhooge, I., & Keppler, H. (2021). Central auditory processing and listening effort in normal-hearing children: A pilot study. *International Journal of Audiology, 60*(10), 739–746. https://doi.org/10.1080/14992027.2021.1877365

Dawes, P., & Bishop, D. V. (2010). Psychometric profile of children with auditory processing disorder and children with dyslexia. *Archives of Disease in Childhood, 95*(6), 432–436. https://doi.org/10.1136/adc.2009.170118

DeBonis, D. A. (2015). It is time to rethink central auditory processing disorder protocols for school-aged children. *American Journal of Audiology, 24*(2), 124–136. https://doi.org/10.1044/2015_aja-14-0037

DeBonis, D. A., & Moncrieff, D. (2008). Auditory processing disorders: An update for speech-language pathologists. *American Journal of Speech-Language Pathology, 17*(1), 4–18. https://doi.org/10.1044/1058-0360(2008/002)

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications.* Sage Publications.

de Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide.* Cambridge University Press. https://doi.org/10.1017/CBO9780511996214

Domitz, D. M., & Schow, R. L. (2000). A new CAPD battery—Multiple auditory processing assessment. *American Journal of Audiology, 9*(2), 101–111. https://doi.org/10.1044/1059-0889(2000/012)

Eggermont, J. J. (2015). *Auditory temporal processing and its disorders.* OUP Oxford.

Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders, 40*(4), 481–492. https://doi.org/10.1044/jshd.4004.481

Erden, G., Kurdoğlu, F., & Uslu, R. (2002). Ilköğretim okullarina devam eden Türk çocuklarinin sinif düzeylerine göre okuma hizi ve yazim hatalari normlarinin geliştirilmesi [Development of reading speed and spelling error norms according to the grade levels of Turkish children attending primary schools]. *Türk Psikiyatri Dergisi, 13*(1), 5–13.

Estabrooks, W., MacIver-Lux, K., & Rhoades, E. A. (2016). *Auditory-verbal therapy: For young children with hearing loss and their families, and the practitioners who guide them.* Plural Publishing.

Evans, S., McGettigan, C., Agnew, Z. K., Rosen, S., & Scott, S. K. (2016). Getting the cocktail party started: Masking effects in speech perception. *Journal of Cognitive Neuroscience, 28*(3), 483–500. https://doi.org/10.1162/jocn_a_00913

Ferenczy, M., Pottas, L., & Soer, M. (2022). Speech perception in noise in children with learning difficulties: A scoping review. *International Journal of Pediatric Otorhinolaryngology, 156*, Article 111101. https://doi.org/10.1016/j.ijporl.2022.111101

Geffner, D., & Goldman, R. (2010). *ASA: Auditory Skills Assessment.* PsychCorp.

Geffner, D., & Ross-Swain, D. (2018). *Auditory processing disorders: Assessment, management, and treatment.* Plural Publishing.

Golding, M., Carter, N., Mitchell, P., & Hood, L. (2004). Prevalence of central auditory processing (CAP) abnormality in an older Australian population: The Blue Mountains hearing study. *Journal of the American Academy of Audiology, 15*(9), 633–642. https://doi.org/10.3766/jaaa.15.9.4

Goldman, R. (2015). *Goldman-Fristoe Test of Articulation–Third edition (GFTA-3)*. AGS.

Hämäläinen, J. A., Salminen, H. K., & Leppänen, P. H. (2013). Basic auditory processing deficits in dyslexia: Systematic review of the behavioral and event-related potential/field evidence. *Journal of Learning Disabilities, 46*(5), 413–427. https://doi.org/10.1177/0022219411436213

Holden, J., Francisco, E., Lensch, R., Tommerdahl, A., Kirsch, B., Zai, L., Dennis, R., & Tommerdahl, M. (2019). Accuracy of different modalities of reaction time testing: Implications for online cognitive assessment tools. *BioRxIV*, Article 726364. https://doi.org/10.1101/726364

Holube, I. (2011). Speech intelligibility in fluctuating maskers. *Proceedings of the International Symposium on Auditory and Audiological Research, 3*, 57–64. https://proceedings.isaar.eu/index.php/isaarproc/article/view/2011-06

Houben, R., van Doorn-Bierman, M., & Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *International Journal of Audiology, 52*(11), 753–761. https://doi.org/10.3109/14992027.2013.832415

Jain, A., Bansal, R., Kumar, A., & Singh, K. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied & Basic Medical Research, 5*(2), 124–127. https://doi.org/10.4103/2229-516x.157168

Jerger, J., & Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorders in school-aged children. *Journal of the American Academy of Audiology, 11*(09), 467–474. https://doi.org/10.1055/s-0042-1748136

Johnston, M., Dixon, D., Hart, J., Glidewell, L., Schröder, C., & Pollard, B. (2014). Discriminant content validity: A quantitative methodology for assessing content of theory-based measures, with illustrative applications. *British Journal of Health Psychology, 19*(2), 240–257. https://doi.org/10.1111/bjhp.12095

Katz, J. (2016). The Buffalo CAPD Model: The importance of phonemes in evaluation and remediation. *Journal of Phonetics & Audiology, 2*(1), Article 111. https://doi.org/10.4172/2471-9455.1000111

Keith, R. W. (2000). Development and standardization of SCAN-C test for auditory processing disorders in children. *Journal of the American Academy of Audiology, 11*(08), 438–445. https://doi.org/10.1055/s-0042-1748131

King, W. M., Lombardino, L. J., Crandell, C. C., & Leonard, C. M. (2003). Comorbid auditory processing disorder in developmental dyslexia. *Ear and Hearing, 24*(5), 448–456. https://doi.org/10.1097/01.aud.0000090437.10978.1a

Kiziltan, E., Barut, C., & Gelir, E. (2006). A high-precision, low cost system for evaluating finger-tapping tasks. *International Journal of Neuroscience, 116*(12), 1471–1480. https://doi.org/10.1080/00207450500514318

Kobayashi, M., Morimoto, M., Sato, H., & Sato, H. (2007). Optimum speech level to minimize listening difficulty in public spaces. *The Journal of the Acoustical Society of America, 121*(1), 251–256. https://doi.org/10.1121/1.2382499

Koiek, S., Pourbakht, A., Mahdavi, M. E., & Tahaei, A. A. (2018). Acceptable noise level in learning disordered children. *Auditory and Vestibular Research, 27*(2), 86–92. https://avr.tums.ac.ir/index.php/avr/article/view/689

Kreisman, N., John, A., Kreisman, B., Hall, J., & Crandell, C. (2012). Psychosocial status of children with auditory processing disorder. *Journal of the American Academy of Audiology, 23*(3), 222–233. https://doi.org/10.3766/jaaa.23.3.8

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Leccese, F., Tuoni, G., Salvadori, G., & Rocca, M. (2015). An analytical model to evaluate the cocktail party effect in restaurant dining rooms: A case study. *Applied Acoustics, 100*, 87–94. https://doi.org/10.1016/j.apacoust.2015.06.012

Libben, G., Gagné, C. L., & Dressler, W. U. (2020). The representation and processing of compounds words. In V. Pirrelli, I. Plag, & W. U. Dressler (Eds.), *Word knowledge and word usage* (pp. 336–352). De Gruyter Mouton. https://doi.org/10.1515/9783110440577-009

Martin, N., & Brownell, R. (2005). *TAP-3: A language processing skills assessment.* WPS Publishing.

Martin, N., Brownell, R., & Hamaguchi, P. (2018). *TAP-4: A language processing skills assessment.* WPS Publishing.

Masters, M. G., Stecker, N. A., & Katz, J. (1998). *Central auditory processing disorders: Mostly management.* Pearson.

Meinzen-Derr, J., Wiley, S., Creighton, J., & Choo, D. (2007). Auditory skills checklist: Clinical tool for monitoring functional auditory skill development in young children with cochlear implants. *Annals of Otology, Rhinology & Laryngology, 116*(11), 812–818. https://doi.org/10.1177/000348940711601104

Meister, H., Rählmann, S., Lemke, U., & Besser, J. (2018). Verbal response times as a potential indicator of cognitive load during conventional speech audiometry with matrix sentences. *Trends in Hearing, 22*, Article 233121651879325. https://doi.org/10.1177/2331216518793255

Moore, D. R. (2007). Auditory processing disorders: Acquisition and treatment. *Journal of Communication Disorders, 40*(4), 295–304. https://doi.org/10.1016/j.jcomdis.2007.03.005

Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., & Ferguson, M. A. (2011). Development of auditory processing in 6- to 11-yr-old children. *Ear and Hearing, 32*(3), 269–285. https://doi.org/10.1097/AUD.0b013e318201c468

Musiek, F. E., & Chermak, G. D. (2007). *Handbook of central auditory processing disorder.* Plural Publishing.

Palana, J., Schwartz, S., & Tager-Flusberg, H. (2022). Evaluating the use of cortical entrainment to measure atypical speech processing: A systematic review. *Neuroscience & Biobehavioral Reviews, 133,* Article 104506. https://doi.org/10.1016/j.neubiorev.2021.12.029

Reynolds, C., Voress, J., & Pearson, N. (2005). *Test of Auditory Processing Skills-III* (3rd ed.). Proedinc Publishing.

Richard, G. J. (2017). *The source – Processing disorders* (2nd ed.). Pro-Ed.

Richard, G. J., & Hanner, M. A. (2005). *Language Processing Test, elementary.* Pro-Ed.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., & Pichora-Fuller, M. K. (2013). The ease of language understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7,* Article 31. https://doi.org/10.3389%2Ffnsys.2013.00031

Schow, R. L., Dillon, H., Hillam, J., Whitaker, M. M., & Seikel, J. A. (2021). Factor analysis on Multiple Auditory Processing Assessment–2 and Listening in Spatialized Noise–Sentences Test in children. *American Journal of Audiology, 30*(2), 433–442. https://doi.org/10.1044/2021_aja-20-00158

Sharma, M., Purdy, S. C., & Kelly, A. S. (2009). Comorbidity of auditory processing, language, and reading disorders. *Journal of Speech, Language, and Hearing Research, 52*(3), 706–722. https://doi.org/10.1044/1092-4388(2008/07-0226)

Smoski, W., Brunt, M. A., & Tannahill, J. C. (1998). *Children's Auditory Performance Scale (CHAPS)*. Educational Audiology Association.

Steadman, M. A., & Sumner, C. J. (2018). Changes in neuronal representations of consonants in the ascending auditory system and their role in speech recognition. *Frontiers in NeuroScience, 12*, Article 671. https://doi.org/10.3389/fnins.2018.00671

Stephan (n.d.). *Restaurant ambiance* [Audio track]. Retrieved February 20, 2021, from https://soundbible.com/1664-Restaurant-Ambiance.html

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (4th ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199231881.001.0001

Tomlin, D., Dillon, H., Sharma, M., & Rance, G. (2015). The impact of auditory processing and cognitive abilities in children. *Ear and Hearing, 36*(5), 527–542. https://doi.org/10.1097/aud.0000000000000172

Ueda, M., & Tanaka, T. (2020). *What is this sound in decibel? –We tried to determine students' comprehension of sound levels, online version* [Paper presentation]. 17th International Conference on Cognition and Exploratory Learning in Digital Age, Japan. https://eric.ed.gov/?id=ED626758

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (1999). *Comprehensive test of phonological processing: CTOPP.* Pro-Ed.

Warrier, C. M., Johnson, K. L., Hayes, E. A., Nicol, T., & Kraus, N. (2004). Learning impaired children exhibit timing deficits and training-related improvements in auditory cortical responses to speech in noise. *Experimental Brain Research, 157*, 431–441. https://doi.org/10.1007/s00221-004-1857-6

Webster, R. E. (2009). *Test of Information Processing Skills.* ATP Assessments.

Wilson, W. J., & Arnott, W. (2013). Using different criteria to diagnose (central) auditory processing disorder: How big a difference does it make? *Journal of Speech, Language, and Hearing Research, 56*(1), 63–70. https://doi.org/10.1044/1092-4388(2012/11-0352)

Yathiraj, A., & Maggu, A. R. (2013). Screening Test for Auditory Processing (STAP): A preliminary report. *Journal of the American Academy of Audiology, 24*(09), 867–878. https://doi.org/10.3766/jaaa.24.9.10

Yathiraj, A., & Vanaja, C. S. (2018). Criteria to classify children as having auditory processing disorders. *American Journal of Audiology, 27*(2), 173–183 https://doi.org/10.1044/2018_aja-17-0091

Zokoll, M. A., Fidan, D., Türkyılmaz, D., Hochmuth, S., Ergenç, İ., Sennaroğlu, G., & Kollmeier, B. (2015). Development and evaluation of the Turkish matrix sentence test. *International Journal of Audiology, 54*(sup2), 51–61. https://doi.org/10.3109/14992027.2015.1074735

Authors' Note

Correspondence concerning this article should be addressed to Hale Hancer, Department of Audiology, Faculty of Health Sciences, Baskent University, Baglica Campus, Ankara, 06810, TÜRKIYE. Email: **halehancer@gmail.com**

Acknowledgments

This research was supported by the Scientific and Technological Research Council of Turkey under the 1002 Quick Support Program (project number 121S261).

Disclosures

The authors declare no conflicts of interest, financial or otherwise. This manuscript was produced from a doctoral thesis titled "Development, Validity and Reliability of Auditory and Speech Performance Test for Children" (Ankara University, Institute of Health Sciences, 2023) by Hale Hancer.