Development and Standardization of SCAN-A: Test of Auditory Processing Disorders in Adolescents and Adults

Robert W. Keith*

Abstract

This paper reports on the development and standardization of SCAN-A: Test of Auditory Processing Disorders in Adolescents and Adults. The standardization version of SCAN-A included six subtests: two filtered words (FW) subtests, two auditory figure-ground (AFG) subtests, a competing words (CW) subtest, and a competing sentences (CS) subtest. Studies included development of normative results on subjects at different ages, determination of subtests to retain in the final version, investigation of test–retest reliability, and a concurrent validity study comparing the results of SCAN-A to SCAN scores obtained on the same subjects. The final version of SCAN-A contains four subtests that take 20 minutes to administer. Raw scores can be converted to standard scores and percentile ranks. Cut-off scores are suggested for normal, questionable, and abnormal performance.

Key Words: Auditory figure ground, central auditory processing disorders, competing sentence test, dichotic speech tests, filtered words, sensitized speech tests

This paper reports on the development and standardization of SCAN-A: Test of Auditory Processing Disorders in Adolescents and Adults. In 1986, Keith published SCAN: A Screening Test for Auditory Processing Disorders. SCAN was designed to be used with children and is normed from ages 3 to 11 years. The purposes of SCAN are (1) to determine possible disorders of central nervous system function, (2) to identify auditory processing problems, and (3) to identify children who may benefit from remediation. Since its publication, the test has received widespread recognition (Musiek et al., 1990; Cherry, 1992) and is used in hundreds of school districts and audiology clinics. While SCAN has met some of audiologists’ needs for central auditory test materials, there continues to be a need for well-standardized central auditory tests for persons over 11 years of age. Such tests should be readily available as a quality recording. For example, Noffsinger (1992) noted that materials used in auditory perceptual assessment are divergent in quality and complexity, and that there is a “tremendous need for standardization of central auditory tasks that would be clinically useful” (p. 10). In addition, the ASHA Ad Hoc Committee on Central Auditory Processing (ASHA, 1991) recommended as a future need “the development of ... measures for identification and assessment of central auditory processing disorders” (p. 10). This would include the development of “batteries of behavioral measures necessary for identification and assessment of central auditory processing disorders” (p. 10). In partial response to that need, Keith (1994) recently modified SCAN for use with older subjects. The name of the new test is SCAN-A: Test for Auditory Processing Disorders in Adolescents and Adults.

SCAN-A is an upward extension of SCAN (Keith, 1986). It is a battery of central auditory tests designed for both adolescents and adults. Its purpose is to assist in determining whether central auditory processing difficulties contribute to an adolescent’s academic difficulties. SCAN-A can be used with adolescents and adults with neurologic disorders and learning disabilities. The purposes are to enable a clinician to identify the social and vocational needs of individuals and provide information about “functional disorders of communication” (Bergman et al., 1987). Documentation of auditory processing disorders also assists employers in making
reasonable accommodations for optimizing the workplace in compliance with the Americans with Disabilities Act, 1991.

In keeping with the substantial need for documentation of emerging central auditory test batteries, the purposes of this article are to describe SCAN-A, its design, and results of standardization studies.

METHOD

The standardization version of SCAN-A included six subtests: two filtered words (FW) subtests, two auditory figure-ground (AFG) subtests, and a competing words (CW) and a competing sentences (CS) subtest. The stimuli for the FW, AFG, and CW subtests were monosyllabic words taken from the Spache readability word list (Spache, 1953) and the Kindergarten Phonetically Balanced word lists (Haskins, 1949). The competing sentences were modified from the Bamford-Kowal-Bench Standard Sentence Lists (Bench and Bamford, 1979). For both the FW and AFG subtests, 2 practice words followed by 20 test words were presented to each ear. Two practice items also preceded each directed ear listening condition of the competing word and competing sentence subtests.

Filtered Words Subtest

All filtering was done using UREI Model 565 Filter sets. This filter provides a slope of 18 dB/octave with noise floors equivalent to an input signal of -70 dB with a 0.15 percent maximum distortion. Three filter sets, connected in series, were used to achieve a slope of 32 dB/octave. The lists were recorded with low-pass filters at 500 Hz and 750 Hz with a filter roll-off of 32 dB per octave.

Auditory Figure-Ground Subtest

Monosyllabic words were recorded in the presence of multitalker speech babble noise. The word list was recorded with a 0 dB and +4 dB signal-to-noise ratio. No single speaker was identifiable or intelligible in the speech babble noise. The background speech was selected for its uniform intensity, and acoustic peaks that naturally occur during spontaneous speech were electronically eliminated by compression. The final variations in acoustic peaks were +1 dB and -2 dB at the 0-dB level.

Competition Words Subtest

The standardization version of this subtest was identical to the CW subtest in SCAN. The subtest was made up of two lists of 25 monosyllabic word pairs that are presented to the right and left ears with simultaneous onset times. The words were matched within 5 msec for equal duration. Semantically unrelated words were paired. The subtest was preceded by four practice items, with instructions identical to the test items. During the presentation of the first 25 word pairs, the subject repeated both words heard, first saying the word heard in the right ear (right-ear first task or directed right task). During the second 25 word pairs, the subject repeated both words, first saying the word heard in the left ear (left-ear first task or directed left task).

Competing Sentences Subtest

For this subtest, selected sentences were adapted from the Bamford-Kowal-Bench Standard Sentence Lists (Bench and Bamford, 1979). Sentences were modified to reflect typical American vocabulary. Sentence pairs that were unrelated in topic were selected for use. The sentences were recorded with simultaneous onset and with offsets within 150 msec and were matched to have approximately the same rhythm. Sentences were recorded at equal intensities on both channels of the stereo tape cassette for binaural presentation. On the standardization version of the CS subtest, four practice items (two practice items per ear) and 26 sentence pairs (13 pairs for each ear) were presented to subjects. Using a directed ear listening task, subjects repeated sentences presented to the right and ignored the left ear during the first 13 pairs, and repeated sentences heard in the left and ignored the right ear during the second 13 pairs.

Technical Specifications and Recording

A male speaker with clear articulation and midwestern United States dialect was selected to record the words for the audio cassette; the same speaker was previously used to record SCAN. All words and sentences were monitored to 0 with a VU meter as they were read. Monosyllabic stimuli were recorded at 5-second intervals, resulting in approximately 4½-second intervals between stimuli. This rate gave the subject adequate time to respond, without prolonging the subtest unnecessarily.
Tape machines used for recording and tape-to-tape transfers were MCI Model JH110B 2-track 1/4" studio recorders. All recordings were done at a tape speed of 71/2 ips, using 185 nanowebers per meter reference operating level. Dolby™ Type "A" Noise Reduction equipment was used to reduce tape hiss and print-through. Program peak levels were limited/compressed using Valley People's Model 700 Gain Brain equipment.

**Standardization Studies**

Three separate studies completed during standardization of SCAN-A are reported here. Experiment 1 was a basic study of normative results on subjects at different ages that included several purposes: (1) to determine the effects of age of subtest scores, (2) to determine which subtests would be retained in the final version of the test, and (3) to develop normative data. Experiment 2 investigated test–retest reliability, and experiment 3 was a concurrent validity study comparing results of SCAN-A to SCAN scores obtained on the same subjects.

**Description of the Experimental Sample for All Studies**

Subjects were obtained from 21 co-investigator sites located in the four major regions of the country. To participate in the testing, subjects were required to have pure-tone audiometric testing at frequencies of 500, 1000, 2000, and 4000 Hz. Subjects with hearing thresholds poorer than 20 dB HL at any frequency were not included in the sample. Subjects whose hearing test results indicated threshold differences between ears of 10 dB or more at two or more frequencies were also not included. In addition, all subjects were native to American-accented English. Persons with obvious mental retardation or speech or language impairment were not accepted for study.

**Examiners**

The SCAN-A examiners included audiologists in 21 university clinics, private practices, and military hospitals and clinics. Fifty audiologists participated in standardization testing. The subjects' responses were written on the score form by the examining audiologists. Subsequently, scoring and data analysis was done by this author.

Testing was conducted in a room that was quiet and free of distractions, disruptions, and outside noise. The audiotaped signal was presented using either a quality stereo audio cassette player and earphones or with a stereo cassette player through a two-channel audiometer using standard earphones. For accurate scoring, headphones were carefully placed on the subject to be sure that the right and left earphones were on the correct ears. The volume control was adjusted to the subject's most comfortable listening level.

**Instructions**

Instructions to the subtests were recorded on the audio tape, and so all subjects received the same instructions through the headphones. All SCAN-A subtests were imitative speech tasks, requiring the subject to repeat the words or sentences. Subjects were encouraged to guess at responses when they were not sure of the word or sentence. Practice items were provided for each subtest. All subjects were able to complete the test.

**Scoring**

For the FW and AFG subtests, a response was scored as correct if the whole word was accurately repeated. A response was scored as incorrect if the word was omitted or substituted. For the CW subtest, credit was given for each word accurately repeated, even if only one word of a pair was accurately repeated or if the words in the pair were repeated in reverse order. For the CS subtest, credit was given when the sentence was repeated verbatim or when there were minor common dialectical substitutions. Responses were clearly in error when there was a change in the meaning of the sentence or when the listener mixed words from the two ears, that is, when there was a lack of "channel separation."

**Experiment 1: Basic Standardization Study**

The purposes of the basic standardization study were (1) to determine the effects of age on subtest scores, (2) to determine which subtests would be retained in the final version of the test, and (3) to develop normative data. Subjects included 125 persons between the ages of 12 and 50 years. Sixty-four (51.2%) were females and 61 (48.8%) were males. Ninety-six (76.8%) were white, 15 (12%) were African American, 10 (8%) were Hispanic, and 4 (3.2%) were persons of other ethnic background. There were 25 subjects in each of the following age groups: 12–14, 15–18, 19–30, 31–40, and 41–50 years.
**Determination of Final Subtests**

Results of average performance of all subjects for the 500- and 750-Hz low-pass filtered (LPF) words subtests were compared. Results indicated that the total subtest mean and standard deviation of performance for subjects was 35.5 (2.4) for 500 Hz LPF and 33.8 (2.9) for 750 LPF. Based on these comparisons, the FW subtest recorded at 750 Hz was deleted from the final version of the test (Keith and Grant, 1993). Similarly, results of average performance of subjects for the 0 and +4 dB signal-to-noise ratio subtests were compared. Results indicated that the total subtest mean and standard deviation of performance was 36.7 (2.3) at 0 dB S/N and 36.2 (2.5) at +4 dB S/N. Based on these comparisons, the word list with the +4 dB signal-to-noise ratio was omitted from the final version of SCAN-A (Keith and Grant, 1993).

Responses to the CW subtest were analyzed to determine which of the word pairs provided the greatest contribution to the final outcome. Correlations for each item on the CW subtest was calculated against the total ear score and CW total score. Ten word pairs that had low or negative correlations with the ear or total score were identified and removed leaving 30 word pairs (15 per ear) in the final version of SCAN-A. Similarly, responses to the competing sentence subtest were analyzed to determine which of the sentence pairs contributed to the final result. A sentence pair was deleted if there was a high frequency of errors by individuals with normal auditory processing abilities. Twenty pairs of sentences (10 per ear) were retained in the final version of SCAN-A.

**Effect of Age**

Table 1 includes means and standard deviations of SCAN-A subtest and total raw scores.

The results indicate that the mean and standard deviation of results on this test were the same for individuals aged 12 through 50 years. The data confirmed that the auditory system of normally developing individuals is typically mature by age 12 years. For this reason, SCAN-A standard scores were collapsed across age ranges.

**Development of Standard Scores for SCAN-A**

Using the mean and standard deviation of subtest scores, standard scores with a mean of 10 and a standard deviation of 3 were developed. The total test standard score was developed with a mean of 100 with a standard deviation of 15. A standard score is a statistic that represents how far an individual's score deviates from the average score of the test group (the mean). This is a scoring scale common to psychoeducational batteries. It allows examiners to compare performance across similarly standardized tests (e.g., the CELF-R and WISC-R) and to compare a subject's performance on test—retest at different ages. Using the standard scores, examiners can determine normal, questionable, and abnormal performance by setting their own criteria. For example, performance is usually considered normal when performance is −1 SD and above, for subtest standard scores of 7 and above or a total standard score of 85 and above. Questionable performance might include subtest standard scores between 4 and 6 or a total standard score between 70 and 84, which is between −1 and −2 SD (16th and 2.5 percentile). Some examiners may choose to consider a wider range of scores as normal—for example, above 1.7 SD—and consider 1.7 to 2.0 SD a questionable range of performance. The decision about choice of cutoff is determined by the examiner’s willingness to under- or overidentify persons who require follow-up. Disordered central auditory processing

<table>
<thead>
<tr>
<th>Age</th>
<th>n</th>
<th>Filtered Words</th>
<th>Auditory Figure Ground</th>
<th>Competing Words</th>
<th>Competing Sentences</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>12-14</td>
<td>25</td>
<td>34.2</td>
<td>2.7</td>
<td>36.2</td>
<td>2.3</td>
<td>52.0</td>
</tr>
<tr>
<td>15-18</td>
<td>25</td>
<td>35.1</td>
<td>2.7</td>
<td>37.3</td>
<td>1.4</td>
<td>54.6</td>
</tr>
<tr>
<td>19-30</td>
<td>25</td>
<td>35.6</td>
<td>2.1</td>
<td>37.0</td>
<td>2.3</td>
<td>54.7</td>
</tr>
<tr>
<td>31-40</td>
<td>25</td>
<td>34.4</td>
<td>4.1</td>
<td>36.4</td>
<td>2.3</td>
<td>53.3</td>
</tr>
<tr>
<td>41-50</td>
<td>25</td>
<td>36.5</td>
<td>2.4</td>
<td>36.3</td>
<td>2.6</td>
<td>55.0</td>
</tr>
<tr>
<td>Total</td>
<td>125</td>
<td>35.2</td>
<td>3.0</td>
<td>36.7</td>
<td>2.2</td>
<td>53.5</td>
</tr>
</tbody>
</table>

Table 1 Means and Standard Deviations of SCAN-A Subtest and Total Raw Scores
Correlations among the SCAN-A Subtests

The construct validity of SCAN-A was evaluated by examining the intercorrelations among the SCAN-A subtest standard scores. Although the primary purpose of SCAN-A is to measure a general construct called auditory processing, four subtests were selected to provide measures of specific aspects of speech recognition commonly associated with this general construct. Therefore, the relationships among SCAN-A subtests and to the general construct were studied. The SCAN-A intercorrelation coefficients across ages are shown in Table 2.

The correlations range from .004 to .507. The CW subtest is moderately correlated with the CS subtest, suggesting that the two subtests measure different but overlapping aspects of auditory processing. The low correlations between the AFG subtest and the other SCAN-A subtests suggest that the AFG subtest taps a different auditory processing skill than the other subtests.

Ear Advantage Studies

It is important to determine whether a discrepancy in the subject’s right and left ear scores is highly unusual. A clinician may interpret a discrepancy between right and left ear scores to be statistically significant, unless such a discrepancy occurs frequently in the normal adult population. For example, a discrepancy between right and left ear scores may not have clinical significance if 25 percent of the normal population shows the same right versus left ear difference, but a discrepancy manifested by only 2 percent of the normal adult population is a cause for concern. The data presented in Table 3 show the prevalence of right ear advantage and left ear advantage on the CW subtest in the standardization sample.

In the standardization sample, 41 of 125 subjects had a right ear advantage. Of the 41 subjects, 92.7 percent had a right ear advantage of 4 or less. Almost 83 percent had a right ear advantage of 2 or less and 19 subjects (46%) had an ear advantage of 0. There were 65 subjects of 125 in the standardization sample who had a left ear advantage. Of these 65 subjects, 90.7 percent had a left ear advantage of -4 or greater. Therefore, a right ear advantage greater than 4 or a left ear advantage more negative than -4 could be considered clinically significant.

Experiment 2: Test–Retest Reliability

Evidence of reliability is provided by data showing that the SCAN-A scores are homogeneous (internal consistency) and dependable and stable across repeated administrations (test–retest reliability). Thirty-eight subjects participated in the test–retest study. The sample included subjects in three age groups: 19–30, 31–40, and 41–50. The between-test interval ranged from 1 day to 5 months, with a mean of 46 days between the two testings. Means and standard deviations for the test–retest sample are reported in Table 4.

The test–retest reliability coefficient for the total test score was .69, and the standard error
Table 4 SCAN-A Test–Retest Means and Standard Deviations for Subtest and Total Raw Scores

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Test</th>
<th>Retest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>Filtered Words</td>
<td>34.7  2.1</td>
<td>34.9  2.5</td>
</tr>
<tr>
<td>Auditory Figure-Ground</td>
<td>36.1  2.0</td>
<td>36.7  2.2</td>
</tr>
<tr>
<td>Competing Words</td>
<td>55.0  3.1</td>
<td>56.1  2.6</td>
</tr>
<tr>
<td>Competing Sentences</td>
<td>19.7  0.5</td>
<td>19.8  0.5</td>
</tr>
<tr>
<td>Total Score</td>
<td>145.5  5.0</td>
<td>147.5  5.0</td>
</tr>
</tbody>
</table>

N = 38.

The high correlations confirm the homogeneity of both the subtest and the total test scores. The low standard error of measurement (SEM) also reflects the high degree of precision in the test scores, and that the obtained scores differ only slightly from the person’s hypothetical true score. For example, a SEM of plus or minus 1 gives a 68 percent level of confidence that the true score is identified, and a SEM of plus or minus 2 gives a 95 percent level of confidence that the true score has been identified.

Experiment 3: Concurrent Validity

The concurrent validity of SCAN-A was assessed by comparing performance on the first three SCAN-A subtests (FW, AFG, and CS) with the corresponding subtests and the composite total score of SCAN. The sample consisted of 29 subjects (14 males and 15 females) aged 12 to 18 years. Both tests were administered in counterbalanced order. Fifteen subjects took SCAN-A first, and 14 subjects took SCAN first. The between-test time interval ranged from 1 to 30 days, with an average of 6 days. The results are shown in Table 6. Note that scores obtained on SCAN were higher than SCAN-A, presumably because of the “easier” listening conditions (e.g., 1000-Hz versus 500-Hz low-pass FW and +8 versus 0 dB S/N ratio for AFG). The CW subtest was revised on SCAN-A to remove word pairs that had low or negative correlations with the ear or total score and include only word pairs that contributed to the final outcome. Therefore, the SCAN-A CW test may be slightly more difficult than the SCAN version. The correlation between the total test standard score on SCAN-A and the composite standard score on SCAN was .59 after correction for attenuation.
DISCUSSION

In summary, the development and standardization of a central auditory test battery designed for subjects between the ages of 12 and 50 years is presented. The results of the standardization data indicated that:

1. There was no effect of age on the subtest results; thus, final standardization results were collapsed across ages.
2. The optimal condition for the low-pass FW subtest was 500 Hz.
3. The optimal condition for the AFG subtest was 0 dB S/N.
4. The CW test from SCAN works equally well for children and adults.
5. The CS test results were extremely stable among subjects. A number of errors on this subtest may prove to indicate a marked auditory processing disorder.
6. Performance variability among normal subjects was minimal.

It was interesting to note that the mean performance on the 750-Hz LPF and +4 dB S/N ratio tests was slightly poorer than the 500-Hz LPF and 0 dB S/N conditions, respectively. In the standardization tape, the 750 and +4 dB S/N conditions were the last of the six subtests. It is possible that the marginally poorer performance on these subtests was due to an order effect, with fatigue becoming a factor after the subject had gone through approximately 35 minutes of somewhat difficult listening. The data analysis indicated that 500-Hz LPF and 0-dB S/N conditions had some errors, thereby avoiding a ceiling effect, with low performance variability. Given these results, it appears that the test has promise to be effective in separating subjects who are normal from subjects with central auditory processing disorder.

The availability of SCAN-A will allow audiologists to obtain, in a reasonable amount of clinical time, norm-referenced results that can be reliably interpreted. It would appear that SCAN-A has the potential of becoming a useful tool for the assessment of central auditory processing abilities for subjects in the age range over which it was normed. Finally, the diagnosis of a central auditory processing disorder is best accomplished using a team approach to assessment and a variety of measures. SCAN-A is only one of several measures that can be used in the diagnosis of this complex disorder.

REFERENCES


