The SCAN-A in testing for auditory processing disorder in a sample of British adults

Piers Dawes
University of Manchester

Key words: Auditory Processing Disorder (APD), adults and adolescents, test reliability

Acronyms: Auditory Processing Disorder (APD), Specific Language Impairment (SLI), Filtered words (FW), Auditory figure-ground (AFG), Competing words (CW), Competing sentences (CS), Adult Non-word Repetition Test (An-rep)

Piers.dawes@manchester.ac.uk

Audiology and Deafness Research Group
School of Psychological Sciences
University of Manchester
Manchester M13 9PL

The SCAN-A in testing for auditory processing disorder in a sample of British adults

Abstract

Objective The SCAN-A (Keith, 1994) is a test for auditory processing disorders in adolescents aged 11+ and adults developed in the USA. It was previously found that the children's version (the SCAN-C) over diagnoses auditory processing disorder in UK children. This study was conducted to assess the use of the SCAN-A with UK adults. ***Design*** Comparison of UK adult's SCAN-A performance to US-based normative data. ***Study sample*** 31 UK adults aged 19 to 64 years (*M* 24, *sd* 10). ***Results*** The UK sample scored significantly worse on three subtests: the Filtered Words (FW), Competing Words (CW) and Competing Sentences (CS) sections as well as on the Total Score. ***Conclusions*** Applying US norms to UK adult's performance results in a high rate of over-identification of listening difficulties. Alternative UK norms are provided and implications for use of the SCAN-A with UK adolescents and adults are discussed.

Introduction

The SCAN tests (SCAN-C for children aged between 5 and 11:11 years (Keith, 2000) and SCAN-A[1] for adults and adolescents aged 11+ years (Keith, 1994)) are the most widely used tests for APD in the US and the UK (Emanuel, 2002; Hind, 2006). Studies with the SCAN-C and its predecessor the SCAN (Keith, 1986) showed that British school children scored significantly worse than the US norms (Dawes & Bishop, 2007; Marriage, King,

[1] A revised version of the SCAN-A, the SCAN-3A was published in 2009. The major difference of the SCAN-3A is the addition of a compressed sentences test. Additionally, the Auditory-Figure Ground subtest (see text) is now presented at three different signal to noise ratios, rather than one. To address earlier criticisms that the SCAN lacks a temporal processing measure (see text under 'Limitations of the SCAN tests'), a gap detection screening test has been added. Revised performance norms are also provided.

Briggs, & Lutman, 2001), and that this results in a high rate of over-identification of listening difficulties. To date, no study has investigated the use of the SCAN-A with UK adults that we are aware of. The current study was conducted to examine whether the adult version of the SCAN suffers from a similar problem as the SCAN-C, and if so, how this might be countered.

The SCAN-A is in the same format as the SCAN-C. Stimuli are recorded on CD and presented via headphones at a comfortable listening level. The SCAN-A may be administered either in quiet room or audiometric conditions with a total testing time of 20 minutes. The SCAN-A provides scores for performance in each of four subtests as well as an overall score. US population-based norms are provided, and performance scores may be transformed into standard scores using these norms. The four subtests comprise: 1. Filtered Words (FW) single monosyllabic words that has been low-pass filtered at 750 Hz to reduce intelligibility; 2. Auditory Figure-Ground (AFG) single monosyllabic words presented against multi-talker babble at +4 dB SNR; 3. Competing Words (CW) two monosyllabic words presented simultaneously, one word to each ear; 3. Competing Sentences (CS) pairs of sentences presented simultaneously, one sentence to each ear. The test-taker is asked to repeat the target word or sentence, which may be directed to either the left or the right ear. For AFG and FW there are 20 items per ear with a total possible raw score for each subtest of 40. For CW, there are 30 word pairs, with a total possible score of 60. For CS, there are 10 pairs of sentences with a maximum score of 20 for the subtest. The maximum possible raw score for the SCAN-A is thus 160. Omissions, substitutions or mispronunciations of the target word or sentence are scored incorrect. Stable errors of articulation are scored correct.

For the CW and CS subtests, an ‘ear advantage’ score may be calculated. Ear advantage scores are thought to reflect hemispheric dominance for language, with abnormal scores being symptomatic of neurological disorder or developmental delay. Test-retest reliability coefficient for the Total Test score is reported in the SCAN-A manual as .69.

Limitations of the SCAN tests

As Auditory Processing Disorder (APD) is conceptualized as a condition in which individuals have disproportionate difficulty listening to signals in noise (ASHA, 2005), the SCAN-A seems to have good face validity. However, the SCAN has been criticized for omitting a temporal processing measure (Bellis, 1996). Someone may do well on SCAN tests of degraded, competing or dichotically presented speech, but still have a speech perception problem due to impaired temporal resolution. Some assert that the SCAN should only be used as a screening test for APD (Bellis, 2003; Medwetsky, 2002), although in clinical practice the SCAN is used as a diagnostic instrument. (Emanuel, 2002; Hind, 2006).

The SCAN tests are intended to measure the pre-cognitive, perceptual stage of auditory processing (Keith, 2000) and no understanding at a cognitive level of phonetic differences between speech sounds is required. However, previous studies examining UK children’s performance on the SCAN tests showed that UK children performed significantly worse than the US norms, and that these differences were largely due to mis-hearing the US accent in which the stimuli are recorded (Dawes & Bishop, 2007; Marriage et al., 2001). Woods, Pena and Martin (2004) also found that a higher proportion of Latino-American children scored in the borderline range than might be expected compared to Anglo-American children. These findings are in line with Rosen's (2005) point that perception of phonetic differences depends on language experience. For non-native speakers, perception of certain speech contrasts is especially difficult, for example, perception of the English /r/-/w/ contrast for native speakers of Japanese. Therefore, any test that uses speech stimuli cannot be tapping only a perceptual stage of processing and rather must draw on phonological processing, which is shaped by language experience. This is a difficulty for the validity of speech-based tests of APD in general; if speech-based tests are sensitive to phonological skills, they may also be sensitive to the phonological impairment characteristic of Specific Language Impairment (SLI) and Dyslexia (Bishop, 1997). Additionally, in the case of the SCAN, one might expect that vocabulary and word-retrieval skills would impact on performance of the FW, AFG and CW subtests, while semantic and syntactic skill would impact on the CS subtest. Memory and attentional skills would seem to impact on the entire test, especially the longer sentence stimuli in the CS subtest. The requirement for verbal responding seems to be a possible additional confound with expressive verbal skills. The British Society of Audiology has recently recommended that APD diagnosis be made on the basis of non-speech tests of APD to avoid possible linguistic confounds (British Society of Audiology, 2005). It certainly seems problematic for both clinical and research purposes that a test for APD is sensitive to phonological and linguistic impairment, making it difficult to dis-entangle auditory from other sorts of impairment. In this study, correlations with the Adult Test of Non-word Repetition (An-rep) (Gathercole & Baddeley, 1996) and SCAN-A subtests were examined to discover whether there was a possible association between phonological memory, as measured by the An-rep, and the SCAN-A. However, the primary aim of the study was to discover if the SCAN-A for UK adolescents and adults has similar problems to the SCAN-C for UK children, and if so, how this might be countered.

Method

Ethical approval was obtained from respective Departmental ethics committees of the Universities of Oxford and York.

Subjects

Power analysis suggested that a sample size of 18 would be required to detect a difference in 5 units with a power of 80% using *t*-test, based on published standard deviations for US norms. Thirty-one UK adults, 20 female and 10 male aged 19 to 64 years ($M = 23.9$, $sd = 10.2$) participated in this study. Participants were drawn from the undergraduate, academic and support staff populations of two UK Universities (Undergraduate students $N = 21$, postgraduate students $N = 6$, academic staff members $N = 1$, support staff $N = 3$). Undergraduate students obtained course credit for participation. The level of education attainment for the academic staff member was Phd, while the support staff had all obtained technical college diplomas. All spoke English as a first language. Ethnic origin, based on self-report was 87% White, 10% Asian and 3% Mixed. The 2001 census reports that 92.1% of the UK population is white (http://www.statistics.gov.uk/census2001/ profiles/UK-A.asp). No participant reported any history of or current difficulties with reading or language.

Exclusion criteria were i) hearing thresholds greater than 25 dB at any frequency within 250 Hz to 8000 Hz on pure tone audiometry, ii) asymmetric audiogram, defined as a difference of more than 15 dB in threshold between the ears at any frequency between 250 and 8000 Hz.. Two trained examiners carried out testing after having completed training in administration and scoring. Participant's responses were audio recorded, and for 25% of the sample, responses were re-scored by the second examiner using the audio recording. Inter-rater reliability was high (Cronbach's α ranged from .98 to 1.00 across the four subtests).

Apparatus

Testing was conducted in a quiet room, as recommended in the SCAN-A manual. Hearing screenings were carried out with a Micro Audiometrics Corp DSP screening audiometer with TDH-39P headphones. The SCAN-A was administered using a Dell Latitude notebook computer via Sennheiser HD600 stereo headphones.

The Adult non-word repetition test (An-rep) (Gathercole & Baddeley, 1996) requires test-takers to repeat 28 multi-syllabic nonsense words, such as "brasterer" or "perplisteronk". A score of 1 is awarded for each correct syllable. In order to standardize administration, an audio recording of the An-Rep stimuli (spoken by the author) was used and played to participants from the notebook computer via headphones. The An-rep is designed to assess phonological working memory. Deficits in phonological working memory are characteristic of children with reading and language difficulties (Bishop, 1997), and these difficulties are known to persist into adulthood (Clegg, Hollis, Mawhood, & Rutter, 2005). Due to time limitations, the An-rep was administered to a subset of 20 participants.

Procedure

The SCAN-A was administered and scored as described in the SCAN-A manual. At the start of testing, the presentation volume was adjusted to a comfortable level for each participant. Recorded instructions and practice items then begin each subtest. The words 'Say the word' precede each test item. Responses were scored correct if the response was either an imitation of the US English form or a British English pronunciation of the same word. Stable errors of articulation were scored correct, while substitutions (e.g. 'ball' for 'all') or omissions of the target word were scored incorrect.

Results

UK versus US performance

Figures 1 and 2 show mean raw scores for UK adults compared to US adults. Examination of normality tests revealed a significant departure from normality for CS subtest for the UK sample, with a significant negative skew. Raw data for US norms was unavailable, so similar tests for normality could not be performed. However, examination of performance norms (as discussed below) suggested a negative skew. As raw data for US norms were not available, rank-order based non-parametric alternatives to t-test was not feasible. However, the relatively large sample sizes and similar distributions of the two groups suggest that t-test would be sufficiently robust to departures from normality in this case, according to guidelines by Heeren and D'Agostino (1987).

UK adults scored significantly more poorly than the US adults for FW, CW and CS subtests as well as for the Total score. Applying US norms to UK adult performance therefore results in a higher than expected number of clinical cases. The SCAN-A uses a normally distributed population-based definition of clinically abnormal levels of performance, with ‘questionable’ performance between -1 and -2 standard deviations, and ‘disordered’ performance below -2 standard deviations. This should result in around 12% and 2% of persons classified as ‘questionable’ and ‘disordered’ respectively, if administered to a representative population-based sample. Cut-off scores for classification for clinical categories ‘normal’, ‘questionable’ and ‘disordered’ are provided in the SCAN-A manual for each subtest and for the Total Score. Applying these cut-offs to the UK sample in this study would yield 23% (7 cases) in the ‘questionable’ range and 16% (5 cases) within the disordered range. A difference in mean performance seems to be one issue with using US norms with UK adults, although a second problem is apparent on examination of normative tables provided in the SCAN-A manual. Distribution of performance scores is highly negatively skewed; most people score close to 100% correct on each subtest. Performance norms based on a Gaussian (normal) distribution have been applied using a linear Z transform. However, because the distribution is not symmetrical, the effect of this seems to be to increase the number of false positives within the clinical range, especially at lower levels of performance. In practice, a very small number of errors could thus result in a very poor normalized score. As the distribution of scores for the US and the UK samples is similar, one could adjust for UK performance by adding a compensatory constant to UK scores prior to comparison with US norms. Alternatively, one might generate a list of acceptable alternative responses for items that were prone to error in UK administration. However, an error analysis (reported below) suggested that while mis-hearing of pronunciation may underlie performance differences between the samples, there were few consistencies in the alternative responses made by test-takers. In any case, neither solution would address the difficulty with generating norms based on a highly skewed, non-normal distribution. A method of addressing both difficulties is to use the UK adult performance in this sample, and model the distribution using a closer-fitting alternative to the normal distribution. Examination of P-P plots suggested that an inverse gamma distribution might fit the data more closely than a normal distribution. Based on the population parameters estimated in this sample of UK adults, cut-offs for the lowest 16% (below -1 SD; 'questionable') and the lowest 2% (below -2SD; 'disordered') were generated using an inverse gamma distribution and are detailed for each subtest in table 1. These cut off scores yielded 10% (3 cases) and 3% (1 case) of 'questionable' and 'disordered' performance, respectively for the UK sample based on Total Score.

TABLE 1 HERE

Error analysis

UK adults' performance of the FW subtest was significantly poorer than the US norms. An error analysis for items with greater than 25% error rate was conducted, and the results are shown in Table 2. The most common alternative responses (those reported over 20% of the time) are also shown. As with UK children, some errors appear to be due to misinterpretation of accent, for example, 'arm' for 'on'. This is in line with the conclusions of earlier research with UK children, which concluded that the SCAN was not purely imitative in that test-takers relied on speech recognition mechanisms to recognize target words (Dawes & Bishop, 2007; Marriage et al., 2001). Where there was a mis-match between the stimulus and the stored representation, UK homophones were typically substituted for the US accented word.

TABLE 2 HERE

Correlations between SCAN and Adult Non-word Repetition Test

There were no significant correlations between the FW, AFG and CW subtests and the An-rep. There was a significant correlation between CS and An-Rep (r = .65, p < .01).

Discussion

The SCAN is a popular test thanks to easy availability, high quality CD recorded stimuli, easy administration and population-based performance norms. However, earlier research showed that the with children's version of the SCAN, the SCAN-C, UK children scored significantly more poorly – close to one standard deviation more poorly – than the US norms. These differences were largely due to accent effects. It was concluded that the SCAN-C over-identifies APD in UK children. The current study suggests that the adult version, the SCAN-A suffers from similar problems.

This paper provides revised performance norms based on the performance of a sample of UK adults, though because this University-based sample likely contains people with high average general ability that are highly practiced test-takers, these norms may be unrepresentative of the general population. One might expect that performance norms for the general population might be slightly lower. The fact that this sample still scored significantly worse than the US norms provided by the SCAN-A makes this finding all the more striking.

A general implication is that speech-based tests of auditory processing are flawed because they are not simply tests of repetition; test-takers rely on phonological processing and mapping of inputs to phonological representations of sounds that are built up through experience. Speech-based tests are therefore liable to be biased against those who are unfamiliar with the accent of the speaker in the stimuli recordings and those who have impaired phonological processing, as is characteristic of those with language or reading difficulties. Note that even if one were to re-record stimuli with a British accent, a potential bias would still exist against those who speak regional accents around the UK, those who speak English as a second language and those with weak phonological or linguistic ability.

A strong correlation was found between a test of phonological memory (the An-Rep) and the Competing Sentences subtest of the SCAN. This is understandable as the sentence content of this subtest places more demands on memory than the single syllable words in the other three SCAN subtests. An implication is that speech-based tests, especially those that use longer sentence-based stimuli may be sensitive to phonological memory impairments, which are characteristic of reading and language difficulties (Bishop, 1997).

In summary, as APD is conceptualised as an impairment in the pre-cognitive, perceptual stage of auditory processing (Keith, 2000), it does seem problematic to use speech-based tests generally to identify APD. It is preferable to adopt the BSA recommendation that

APD be diagnosed using non-speech tests in order to avoid confound with phonological or linguistic impairment. If speech-based tests such as the SCAN-A must be used, they should be used conservatively – i.e. by not relying on the results of SCAN-A alone for APD diagnosis, and in association with appropriate performance norms.

Acknowledgements

This research was supported in part by a grant from Deafness Research UK. Thank you to Ms Olga Alexatou for assistance in data collection.

Declaration of interest

There are no conflicts of interest.

References

American Speech-Language Hearing Association. (2005). (Central) auditory processing disorders [Electronic Version], 1-20. Retrieved 2005 from http://www.asha.org/docs/html/TR2005-00043.html.

Bellis, T. J. (1996). *Assessment and management of central auditory processing disorders in the educational setting: from science to practice.* San Diego: Singular Publishing Group.

Bellis, T. J. (2003). *Assessment and management of central auditory processing disorders in the educational setting: From science to practice* (2nd ed.). Clifton Park, NY: Delmar Learning.

Bishop, D. V. M. (1997). *Uncommon understanding*. Hove: Psychology Press.

British Society of Audiology. (2005). Working definition of APD. Retrieved 20th September 2005, from http://www.thebsa.org.uk/apd/Home.htm#working%20def;2005

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders - a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry*, *46*(2), 128-149.

Dawes, P., & Bishop, D. V. M. (2007). The SCAN-C in testing for Auditory Processing Disorder in a sample of British children. *International Journal of Audiology, 46*(12), 780-786.

Emanuel, D. C. (2002). The auditory processing battery: Survey of common practices. *Journal of the American Academy of Audiology, 13*, 93-117.

Gathercole, S. E., & Baddeley, A. D. (1996). *Adult test of nonword repetition*. Bristol: University of Bristol.

Heeren, T., & D'Agostino, R. (1987). Robustness of the two independent samples t-test when applied to ordinal scaled data. *Statistics in Medicine*, *6*(1), 79-90.

Hind, S. (2006). Survey of care pathway for auditory processing disorder. *Audiological Medicine*, *4*(1), 12-24.

Keith, R. W. (1994). *SCAN - A: A Test for Auditory Processing Disorders in Adolescence and Adults*. San Antonio: Psychological Corporation.

Keith, R. W. (Ed.). (1986). *SCAN: A Screening Test for Auditory Processing Disorders in Children*. San Antonio: Psychological Corporation.

Keith, R. W. (Ed.). (2000). *SCAN-C Test for Auditory Processing Disorders in Children - Revised*. San Antonio: Psychological Corporation.

Keith, R. (2009). *SCAN-3A*. Oxford: Pearson Assessment.

Marriage, J., King, J., Briggs, J., & Lutman, M. E. (2001). The reliability of the SCAN test: Results from a primary school population in the UK. *British Journal of Audiology, 35*, 199-208.

Medwetsky, L. (2002). Central auditory processing testing: A battery approach. In J. Katz (Ed.), *Handbook of Clinical Audiology* (5th ed.). Philadelphia: Lippincott, Williams & Wilkins.

Rosen, S. (2005). "A riddle wrapped in a mystery inside an enigma": Defining central auditory processing disorder. *American Journal of Audiology, 14*, 139-142.

Woods, A. G., Pena, E. D., & Martin, F. N. (2004). Exploring possible sociocultural bias on the SCAN-C. *American Journal of Audiology, 13*(2), 173-184.

Table 1. Cut-off scores for UK performance norms

Subtest	Cut-off score
Filtered Words	
<1 SD 'Questionable'	29
<2 SD 'Disordered'	24
Auditory Figure Ground	
<1 SD 'Questionable'	34
<2 SD 'Disordered'	31
Competing Words	
<1 SD 'Questionable'	47
<2 SD 'Disordered'	40
Competing sentences	
<1 SD 'Questionable'	16
<2 SD 'Disordered'	14
Total Score	
<1 SD 'Questionable'	130
<2 SD 'Disordered'	122

Table 2. Items with error rates over 50% in the FW subtest

RIGHT EAR				
did	25%	dead 60%		
need	52%	None consistent		
own	25%	on 60%		
leave	43%	lead 56%	read 22%	
on	43%	arm 44%	ant 33%	aren't 22%
find	33%	fine 100%		
true	76%	crew 38%		
ship	100%	sip 57%		
LEFT EAR				
grew	33%	gruel 29%		
mouth	43%	mouse 56%		
such	100%	None consistent		
card	25%	hard 60%		
way	29%	wade 50%		
five	25%	vie 40%		
box	52%	back 91%		
hit	90%	hip 78%		
is	57%	give 25%		
sing	95%	thing 75%		

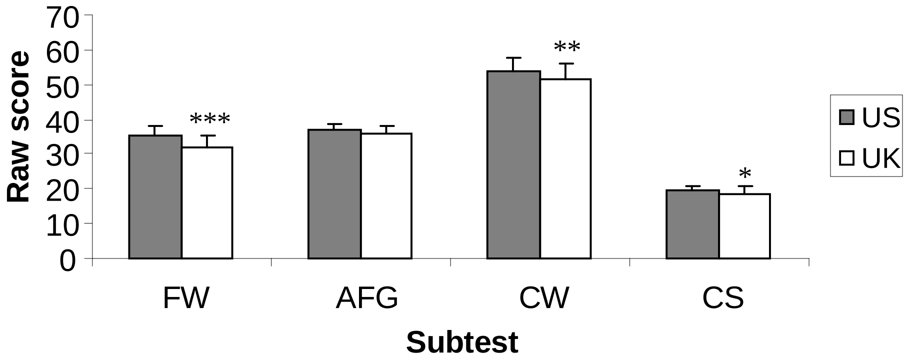

Figure 1. UK versus US adults performance by subtest

Student's t-test comparing US/UK differences: *p<0.05; **p<0.01, ***p<0.001

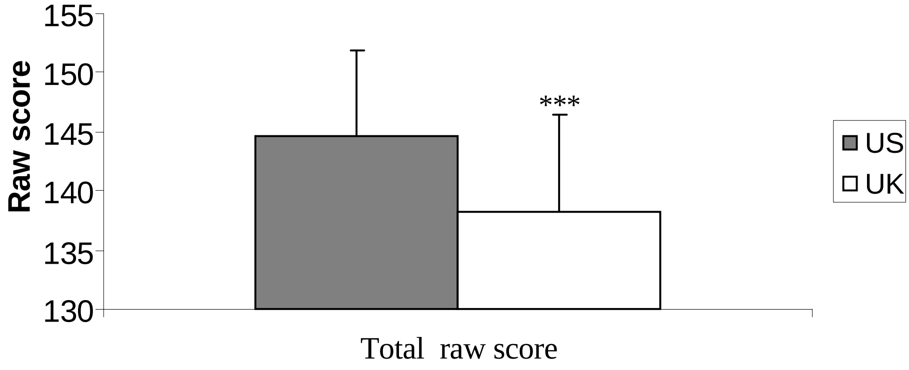

Figure 2. UK versus US adults performance: Total raw score

Student’s t-test comparing US/UK differences: ***p<0.001