Impact of different cutoff criteria on rate of (central) auditory processing disorders diagnosis using the central test battery

Mohsin Ahmed Shaikh,1 Lisa Fox-Thomas,2 Denise Tucker2

1Department of Audiology and Speech-Language Pathology, Bloomsburg University of Pennsylvania; 2Department of Communication Sciences and Disorders, University of North Carolina at Greensboro, USA

Abstract

The purpose of this study was to quantify how the use of two different cutoff criteria affects the test failure rate and potential diagnosis of central auditory processing disorder (C)APD in a sample of children subjected to central auditory processing (C)AP assessment.

Test failure rates for the central test battery (CTB) using two different cutoff criteria (1 and 2 SDs below the mean) were measured retrospectively for 98 children who completed (C)AP assessment. The rates of potential (C)APD diagnosis ranged from 86.8% [when a 1 standard deviation (SD) cutoff was used] to 66.2% (when a 2 SD cutoff was used). The current use of two different cutoffs for the CTB has a large impact on the diagnostic rate for (C)APD. These findings have clinical implications for the diagnosis of (C)APD due to the widespread use of the CTB in the United States for the assessment of (C)APD in children. Thus, it is important to create awareness among audiologists that use of the 2 SDs cutoff criterion is recommended for reducing false positives (error).

Introduction

Identifying children as having central auditory processing disorder (C)APD can be a challenge for educators and audiologists due, in part, to lack of a “gold standard” for differential diagnosis. Although standardized central auditory processing (C)AP tests are routinely used in the United States, there is no universally-accepted (C)APD test battery. Similarly, the choice of cutoff score for “significant” (abnormal) test findings directly impacts diagnosis. Currently, there is a lack of agreement among researchers and clinicians about the cutoff scores that should be used. Some researchers recommend 2 SDs below the mean as the cutoff score that should be used to classify children as pass or fail on (C)AP tests.1-4 However, others support using 1 SD below the mean as the cutoff score.5 Because two different cutoff scores are used, clinicians identify children with (C)APD at different rates. As a result, the proportion of children with lower scores on (C)AP tests who actually have (C)APD remains uncertain.

To diagnose (C)APD, audiologists must conduct a test battery that assesses various auditory processes, such as auditory discrimination, temporal aspects of audition, localization and lateralization, auditory pattern recognition, auditory performance in the presence of competing acoustic stimuli and auditory performance with degraded acoustic signals.2 The clinical guidelines developed by the American Academy of Audiology (AAA) and the American Speech-Language-Hearing Association (ASHA) provide detailed recommendations about the test principles, types of (C)AP tests, and diagnostic criteria.1,2 However, audiologists are free to select a test battery based on the referral complaint, the patient’s age and motivation level, and other cognitive linguistic factors.

Numerous tests are currently available that clinicians may use to assess (C)AP.1,2 The Central Test Battery (CTB) is one of the most widely used test batteries in the United States to diagnose (C)APD in adults and children.7,9 The CTB, which was developed by Jack Katz,5 consists of three widely used tests: the Staggered Spondaic Word (SSW) test, the Phonemic Synthesis (PS) test, and a speech-in-noise (SN) test. The CTB is presented to children on a recorded CD and assesses a variety of auditory processes. The CTB can be used to assess children as young as five years of age. The limitation of the CTB is its lack of any temporal processing tests.

The recommended scoring system for the CTB provides quantitative as well as qualitative metrics. Quantitative scores are computed as number of errors, number of test items correct, and/or percent correct. The qualitative measures provide an additional source of information.
about a participant’s performance on a test (such as information about the response, reversal, delay, perseveration, speed, etc.). The qualitative measurements also allow audiologists to feel confident when making a diagnosis and treatment recommendations. The CTB includes a software tool for data analysis to identify pass/fail status based on the test scores (provided by the clinician).

The CTB currently recommends use of different age norms for children and adults because the central auditory nervous system continues to develop during childhood. While the CTB uses age-specific norms for children 5 – 11 years of age, adult norms are used for children 12 years and above. The CTB also recommends different cut-off criteria for determining pass/fail status on a test. The SD for the test scores of 7-year-old children was reported to be 11% compared to 4% for those of adults. Because children have relatively large SDs even when they are within the same age group, Katz hypothesized that using 2 SDs below the mean as the cutoff would result in too many (C)APD children being overlooked. Therefore, when scoring the CTB, Katz recommends that a 1 SD cutoff be used for younger children (5 – 11 years) and a 2 SD cutoff be used for older children (age 12 years and up) and adults. These two different criteria are built into the CTB software system, which audiologists often use when making clinical diagnoses.

In contrast, the AAA and ASHA do not differentiate between test criteria used for adults and children. They recommend a cutoff criterion of 2 SD below the mean for at least one ear on at least two different behavioral auditory processing tests. If the performance is abnormal on only one test, the cutoff score should be 3 SD below the mean. However, these criteria are not without limitations. First, this assumes a normal distribution. In a distribution that is skewed, performance below the 10th percentile rank on at least one test or performance below the 25th percentile rank on at least two of the behavioral auditory processing tests might be considered abnormal instead. Furthermore, use of the 2 SD criteria is limited by that number of tests included in the battery include specific to the auditory processing deficit. If only one test is included, then the default to a 3 SD cutoff requires the individual to have a profound auditory deficit in that area.

While the ASHA and AAA guidelines recommend a (C)AP test cutoff score for children of 2 SDs below the mean, audiologists continue to use the CTB software that employs a 1 SD cutoff for children. Little research has been conducted regarding the effects of using 1 vs 2 SDs as the cutoff score (to determine failure rates) for the widely-used CTB. Additionally, the impact of the cutoff on test failure rates and the resulting (C)APD diagnostic rate has not yet been reported. It is therefore vital to examine this effect and create awareness among audiologists who use currently the CTB to diagnose (C)APD.

The purpose of this study was to quantify how the two different cutoffs affect the test failure rate and the potential diagnosis of (C)APD in a sample of children subjected to (C)AP assessment. The study further tested Katz’s hypothesis that using a 2 SDs below the mean cutoff compared to a 1 SD below the mean cutoff will fail to identify a significant number of children with (C)APD, particularly a number higher than the expected difference (13.59%) based on a normal distribution. To determine this, to patients’ CTB test scores were classified as pass or fail using both 1 SD below the mean and 2 SDs below the mean criteria. The difference in the number of participants who failed according to the two criteria was then compared with the expected difference based on a normal distribution (13.59%).

**Materials and Methods**

The study was conducted by reviewing medical records. The records of children assessed for (C)AP at the Speech and Hearing Center at the University of North Carolina Greensboro (UNCG) were retrieved and analyzed retrospectively. Permission to access patient files for research was obtained through the UNCG Institutional Review Board (IRB). Demographic data were obtained to create a de-identified (C)APD computer database that included each of the patient’s case history, age, and (C)AP and hearing test results.

**Participants**

Out of the 203 patient files within the UNCG (C)APD computer database, those of 98 children met the following inclusion criteria:

- Normal hearing sensitivity (thresholds <25 dB HL between 250 Hz and 8 kHz, measured at octave intervals);
- Normal middle ear function with compliance and pressure between -100 and +50 daPa;
- No neurological disorders; and
- English as the primary language.

Demographic data and three CTB test scores were collected for these children (F=36, M=62), who had a median of age 9 years (min=7, first quartile=7, third quartile=10, max=11).

**Table 1. Central test battery tests and protocols.**

| CTB test                        | Protocol | Type                                      | Stimuli                          | Score:                           |
|---------------------------------|----------|-------------------------------------------|----------------------------------|----------------------------------|
| Staggered Spondaic Word (SSW)   | Level: 50 dB SL in both ears Type: Dichotic Stimuli: Spondee words Score: Count the incorrect responses (errors) |
| Phonemic Synthesis (PS)         | Level: 50 dB SL in both ears Type: Binaural Stimuli: Strings of discrete phonemes Score: Count the correct responses |
| Speech-in-noise (SN)            | Level: 40-50 dB SL with an SNR of +5 Type: Monaural Stimuli: Monosyllabic words Score: Percent correct score for each ear |

CTB, central test battery; SL, sensation level; SNR, signal to noise ratio.
Data reduction and analysis

The scores for the three CTB tests were retrieved from the patient files for the retrospective analysis. The test score for each of the measured components of the test was included in the analysis. For the SSW test, five components (right non-competing [RNC], right competing [RC], left non-competing [LNC], left competing [LC], and total score) were included. Quantitative and qualitative scores for the PS test and right and left scores for the SN test were included. The test score for each component of the test was compared with age-specific norms. The patient’s CTB test was classified as “fail” if they failed at least one component of the test in at least one ear. The age-specific norms were for these tests that were included in the analysis were obtained from the test manual. Cutoffs were calculated at 1 SD and 2 SDs below the mean. The analysis included the following steps:

- The raw scores were compared to a 1 SD below the mean cutoff to classify children as pass or fail on a test.5
- The raw scores were compared to 2 SDs below the mean cutoff to classify children as pass or fail on a test.1,2
- Out of three tests, the number of tests failed by each child in step 1 was calculated;
- Out of three tests, the number of tests failed by each child in step 2 was calculated;
- For both criteria, children who failed 2 or more tests were classified as (C)APD.

Descriptive and frequency analyses were used to compare the failure rate on each test and the (C)APD diagnostic rate. Chi-square analysis was utilized to compare the difference in the failure rates for the two criteria with the expected difference based on a normal distribution. The level of significance for the chi-square was fixed at 0.05.

Results

The failure rates on the (C)AP tests were analyzed in two ways. First, the failure rate for each of the individual components of the three CTB tests was measured using cutoff scores of 1 SD and 2 SDs below the mean (Table 2). As expected, the failure rate for all components of all tests differed when a cutoff score of 1 SD vs 2 SDs was used. The failure rate was higher (more children failed the test component) for the 1 SD cutoff compared to the 2 SD cutoff. Failure rates on individual components of the SSW ranged from 48.5% to 82.5% when 1 SD was used, whereas failure rates ranged from 39.2% to 70.1% when 2 SDs were used. In the SSW test, the RNC component had the lowest failure rate for both cutoffs. The LC condition had the highest failure rate for both cutoffs, which is consistent with the literature.9 The poor LC performance is suggestive of right ear advantage or left ear weakness commonly observed in children with (C)APD.

As expected, there were also higher failure rates on the PS test when the 1 SD criterion was used. For the quantitative scores, about half (51.5%) of the children failed using the 1 SD cutoff, whereas approximately one-third (36.1%) of the children failed using the 2 SD cutoff. Additionally, the quantitative scores of the patients were better than the qualitative scores. On the SN test, no differences were observed between ears. Almost three-fourths of the children (75.4% for the right and 76.8% for the left) failed using the 1 SD criterion. Using the 2 SD criterion, approximately half of the children still failed the test.

The failure rate on each test was measured using cutoff scores of 1 SD and 2 SDs below the mean. As shown in Table 3, the test failure rate was higher using cutoff scores of 1 SD compared to 2 SDs for each test. Children were classified as having (C)APD if they failed two or more (C)AP tests. As expected, more children were classified as having (C)APD when the 1 SD cutoff score (86.8%) was used compared to when the 2 SD cutoff was used (66.2%). This would equate to an additional 1 in every 5 children being diagnosed with (C)APD using the 1 SD criterion.

The difference in the failure rates obtained using the 1 SD and 2 SD criteria in our (C)AP sample was compared with the expected difference for a normal distribution. In a normally distributed sample, approximately 15.86% of participants fall below a 1 SD below the mean cutoff, and approximately 2.27% of participants fall below the cutoff of

| Test       | Children who fail using 1 SD (%) | Children who fail using 2 SD (%) | Difference (%) | Expected difference (%) | Chi-square (χ², P) |
|------------|----------------------------------|----------------------------------|----------------|--------------------------|-------------------|
| SSW (N=97) | 93.8                             | 81.4                             | 12.4           | 13.59                    | χ²=1.43           |
| PS (N=97)  | 58.8                             | 41.7                             | 17.1           | 13.59                    | P=0.51            |
| SN (N=69)  | 84.1                             | 68.1                             | 16             | 13.59                    |                   |
| Failed at least 2 tests (N=68) | 86.8 | 66.2 | 20.6 | 13.59 |                   |

SD, standard deviation; CTB, central test battery; RNC, right non-competing; RC, right competing; LC, left competing; LNC, left non-competing; SSW, staggered spondaic word; PS, phonemic synthesis; SN, speech in noise.

---

Table 2. Failure rates for components of each of the three central test battery tests using cutoff scores of 1 standard deviation (SD) and 2 SDs below the mean.

| CTB Test | RNC (%) | RC (%) | LC (%) | LNC (%) | Total (%) | PS (N=97) Quantitative (%) | Qualitative (%) | SN (N=69) Right ear (%) | Left ear (%) |
|----------|---------|--------|--------|---------|-----------|---------------------------|----------------|------------------------|-------------|
| SSW (N=97) | 48.5 | 64.9 | 77.3 | 64.9 | 82.5 | 51.5 | 61.1 | 75.4 | 76.8 |

Table 3. Comparison of observed and expected failure rates using cutoff scores of 1 standard deviation (SD) and 2 SDs below the mean.
2 SDs below the mean. Thus, 13.59% (15.86–2.27=13.59) of participants will fall between 1 SD and 2 SDs in a normal distribution. To test Katz’s hypothesis that using a 2 SD criteria will miss too many children with (C)APD, the difference in the test failure rates for the two criteria (1 SD and 2 SD cutoffs) was compared with the expected difference of 13.59%. A chi-square analysis of goodness-of-fit was performed to determine whether the sample of children assessed for (C)APD could be considered part of a normal distribution. The results of the analysis revealed the absence of a difference (2, N=98)=1.43, P=0.51), indicating that the distribution of the current sample was similar to a normal distribution between 1 and 2 SDs.

Discussion

The purpose of this study was to measure the effect of two cutoffs criteria on the test failure rate and the potential diagnosis of (C)APD in a sample of children subjected to (C)AP assessment.

Lack of a consensus in defining a cutoff score for (C)AP tests are evident among audiologists across the country. The CTB recommends using a cutoff of 1 SD for children, whereas the ASHA and AAA recommend using 2 SDs below the mean as the cutoff for (C)AP test scores. This difference in cutoff scores impacts the classification of children as failing the tests as well as the diagnosis of (C)APD. The (C)APD diagnostic rate was 20% higher using the 1 SD cutoff score (86.8%) compared to the 2 SD cutoff score (66.2%). Children who were not classified as having (C)APD in this study may have temporal processing deficits that will not be identified by the CTB. As reported earlier, these children were assessed with other (C)APD tests including temporal processing tests which were not included in this study.

The current study tested Katz’s hypothesis that many children with (C)APD will be missed using the 2 SD criterion. (C)AP assessments for the current sample of children were part of a normal distribution, indicating that changing the cutoff point from 2 SDs to 1 SD leads to the inclusion of more children from a normal distribution. If Katz’s hypothesis had been true, then a distribution would have been observed in which more children (>13.59%) fell between the 1 SD and 2 SDs below the mean cutoffs compared to the corresponding percentage expected for a normal distribution. This type of distribution would provide strong support for the use of 1 SD below the mean as the criterion for failure and the identification of children with (C)APD. Because the results of the current study indicated a normal distribution in this pediatric patient population referred for (C)APD testing, it can be concluded that the use of a 2 SD cutoff will identify children with (C)APD symptoms and may potentially reduce false positives (type I errors). In addition, the higher standard of 2 SDs would increase the confidence of (C)APD diagnosis.

Internationally, in psychology, sociology and the behavioral sciences, 2 SD cutoffs are used to classify children as pass or fail on behavioral tests. A criterion of 2 SDs below the mean was adopted as a standard to differentiate between normal and abnormal performance with greater statistical confidence. Given the findings of this study, it is recommended that audiologists who use the CTB for (C)APD diagnosis consider using 2 SD cutoff scores for children, in compliance with the AAA and ASHA clinical guidelines.

Conclusions

Using different cutoff criteria for (C)AP tests increases the risk of misidentifying children as having (C)APD by increasing the number of false positives (type I errors). (C)AP assessments for the current sample of children were representative of a normal distribution between 1 and 2 SDs. Thus, using a 1 SD below the mean criterion results in the inclusion of more children from the normal distribution and their classification as having (C)APD. As a result, the use of 2 SDs below the mean as a cutoff criterion for children is recommended. The findings of the current study have implications for the clinical diagnosis of (C)APD.

References

1. American Academy of Audiology. American Academy of Audiology clinical practice guidelines: diagnosis, treatment and management of children and adults with central auditory processing disorder; 2010. Available from: http://www.audiology.org/publications-resources/document-library/central-auditory-processing-disorder
2. American Speech-Language-Hearing Association. (Central) auditory processing disorders: technical report; 2005. Available from: http://www.asha.org/policy/tr2005-00043.htm Accessed: April 16, 2016.
3. Chermak GD, Musiek FE, Craig CH. Central auditory processing disorders: new perspectives. San Diego, CA: Singular Publishing Group; 1997.
4. Musiek FE, Chermak GD. Handbook of central auditory processing disorder auditory neuroscience and diagnosis. San Diego, CA: Plural Publishing; 2014.
5. Katz J. Classification of auditory processing disorders. In: Katz J, Stecker NA, Henderson D, eds. Central auditory processing: a transdisciplinary view. St. Louis, MO: Mosby; 1992. pp 81-91.
6. Emanuel DC. The auditory processing battery: survey of common practices. J Am Acad Audiol 2002;13:93-117.
7. Chermak GD, Traynham WA, Seikel JA, Musiek FE. Professional education and assessment practices in central auditory processing. J Am Acad Audiol 1998;9:452-65.
8. Schow RL, Chermak G. Implications from factor analysis for central auditory processing disorders. Am J Audiol 1999;8:137-42.
9. Jutras B, Loubert M, Dupuis JL, et al. Applicability of central auditory processing disorder auditory neuroscience and diagnosis. Am J Audiol 2007;16:100-6.
10. Rawool VW, ed. Auditory processing deficits: assessment and intervention. New York, NY: Thieme; 2015.
11. Higson JM, Haggard MP, Field DL. Validation of parameters for assessing obscure auditory dysfunction-robustness of determinants of OAD status across samples and test methods. Brit J Audiol 1994:28:27-39.
12. Dillon H, Cameron S, Glyde H, et al. An opinion on the assessment of people who may have an auditory processing disorder. J Am Acad Audiol 2012;23:97-105.
13. Jerger J. Clinical experience with impedance audiometry. Arch Otolarngol 1970;92:311-24.