Double Checking the Cross-Check Principle

Robert G. Turner*

Abstract

The cross-check principle was proposed by Jerger and Hayes over 20 years ago. Actually, the cross-check principle consists of a concept and a specific test protocol. The concept, that one test result confirms another test result, is still in use. Their specific protocol is essentially forgotten. The cross-check protocol they described, however, differs from more traditional test protocols. From the design of their cross-check protocol we can extract two unique testing strategies. The first strategy requires agreement between both tests before a decision can be made. This strategy can significantly improve testing performance; however, the “cost” of this strategy is a population of individuals for whom no decision is possible because the two tests disagree. The cross-check protocol uses a third test when the first two tests disagree. Essentially, the cross-check protocol employs an intermediate criterion, the second unusual strategy. Somewhat surprisingly, this intermediate criterion has the ability to simultaneously increase protocol hit rate and decrease protocol false alarm rate relative that of the individual tests in the protocol. The more traditional protocol criteria cannot do this. Each strategy offers some interesting and important advantages and should be considered by audiologists when using multiple tests.

Key Words: Audiology, cross-check principle, false alarm rate, hearing, hit rate, test battery, test correlation, test protocol

Abbreviations: BOA: behavioral observation audiometry; BSER: brain stem evoked response audiometry; FA: false alarm rate; HL: number with hearing loss; HT: hit rate; NH: number with normal hearing; NTD: noise-tone-difference test; UD: Percent of tested individuals for whom no decision is made; VRA: visual reinforcement audiometry

Sumario: El principio del “chequeo cruzado” fue propuesto por Jerge y Hayes hace más de 20 años. De hecho, el principio del “chequeo cruzado” consiste de un concepto y de un protocolo específico de evaluación. El concepto, la idea de que el resultado de una prueba confirma otro resultado de otra prueba, todavía se utiliza. Su protocolo específico, esencialmente, ha sido olvidado. El protocolo de “chequeo cruzado” que ellos describieron, sin embargo, difiere de los protocolos de evaluación más tradicionales. A partir del diseño de dicho protocolo de “chequeo cruzado” podemos extraer dos estrategias singulares de evaluación. La primera estrategia requiere de acuerdo entre dos pruebas antes de poder tomar una decisión. Este enfoque puede mejorar significativamente el rendimiento de la evaluación, aunque el “costo” de esta estrategia es una población de individuos para quienes no es posible tomar una decisión cuando las dos pruebas no están de acuerdo. Esencialmente, el protocolo de “chequeo cruzado” emplea un criterio intermedio, la segunda estrategia inusual. Sorprendentemente, este criterio intermedio tiene la capacidad de incrementar simultáneamente la tasa de aciertos del protocolo, y disminuir la tasa de falsas alarmas, con relación a las correspondientes en las pruebas individuales dentro del protocolo. El criterio del protocolo más tradicional no puede hacer esto. Cada estrategia ofrece algunas ventajas importantes e interesantes, y debería ser considerada por los audiólogos cuando utilicen pruebas múltiples.

Palabras Clave: Audiología, principio del “chequeo-cruzado”, tasa de falsas alarmas, audición, tasa de aciertos, correlación de la prueba, protocolo de la prueba.

*Department of Communication Disorders, Louisiana State University Health Sciences Center, 1900 Gravier St., Box G6-2 New Orleans, LA  70112,  Phone: 504 568-4339, Fax: 504 568-5362, Email: rturner@lsuhsc.edu
J erger and Hayes (1976) proposed the cross-check principle over twenty years ago because of their legitimate concern as to the reliability of the behavioral tests available at that time. Given the poor performance of these tests, the probability of error was high. They recommended that behavioral results be confirmed by a physiological test, immittance or auditory brainstem response audiometry (called BSER by Jerger and Hayes), before making a final decision as to hearing status. Thus, their motivation was not that more tests are always better, but that in this particular situation, there must be some confirmation of behavioral results (Jerger, personal communication).

The cross-check principle actually consists of a concept and a specific test protocol. The concept was explained by Jerger and Hayes. They stated, “...our audiological evaluation of the child does not stop with behavioral test results. We always insist that the behavioral result be confirmed by a cross-check.” Thus, the basic concept is that one test result is used to confirm another test result. As a concept, the cross-check principle has a certain face validity. Requiring agreement among several tests increases confidence in the test results. In addition, disagreement among test results raises a flag that there may be a problem. These characteristics are particularly useful when evaluating a population where testing may be difficult and unreliable. That is why the cross-check principle has been particularly popular in pediatric evaluation.

The cross-check principle is also a specific test protocol as recommended by Jerger and Hayes. In their own words, “We used impedance audiometry to confirm behavioral test results... In most cases, impedance audiometry serves as an effective cross-check of behavioral audiometry. If they disagree, however, BSER audiometry can serve as a further cross-check... We use BSER to cross-check behavioral test results whenever impedance audiometry is noncontributory... We also use BSER to cross-check the results of impedance audiometry when behavioral testing yields no useful information.” Essentially, they proposed a three-test strategy. All individuals receive the first two tests, some type of behavioral test and immittance. The third test, BSER, is administered only when the first two tests disagree. Their cross-check protocol is shown in Figure 1.

**Figure 1** Cross-check protocol recommended by Jerger and Hayes (1976). In their particular protocol, Test 1 is some type of behavioral test; Test 2 is immittance audiometry; Test 3 is auditory brainstem response audiometry (BSER). All individuals receive Test 1 and Test 2. When the results of Test 1 and Test 2 disagree, Test 3 is administered.
Much has changed since Jerger and Hayes proposed the cross-check principle; however, the cross-check principle is still recommended for pediatric and, occasionally, adult testing. Recently, this author was reviewing a book chapter that discussed the use of test batteries for pediatric hearing evaluation and the importance of the cross-check principle. Current guidelines for early hearing detection and intervention state, “Adequate confirmation of an infant's hearing status cannot be obtained from a single test measure. Rather, a test battery is required to cross-check results of both behavioral and physiologic measures” (Finitzo et al., 2000).

What has actually survived is the concept of the cross-check principle, the basic understanding that one test result is used to confirm another test result. What has not survived is the specific test protocol recommended by Jerger and Hayes. This is not surprising since that specific protocol was determined by the cost and availability of tests in the 1970s. Since that time, tests have been modified and new tests, such as otoacoustic emissions, have been introduced. There is no reason to believe that their specific protocol would be optimum today.

Gone unnoticed is the fact that the protocol recommended by Jerger and Hayes differs from more traditional test protocols. Thus, the objective of this paper is to evaluate the particular test protocol described by Jerger and Hayes to determine if this testing strategy offers any advantages over more traditional test protocols. The objective is not to quantitatively evaluate the performance of the cross-check principle as implemented by Jerger and Hayes in the 1970s. The necessary data are probably not available and the need at that time to confirm behavioral results was evident. Before evaluating the cross-check protocol, it is useful to consider more traditional test protocols.

**TRADITIONAL TEST PROTOCOLS**

When constructing a protocol, individual tests are typically combined in parallel or series (Turner, 1988). Parallel means that all tests are given and a criterion is established that specifies how many tests must be positive for the protocol to be positive for the condition of interest, e.g., hearing loss. The most common criteria are strict and loose. Strict means that all individual tests must be positive for the protocol to be positive. Loose means that only one test must be positive for a positive protocol.

A series protocol means that tests are given sequentially and the result on one test determines if the next test is given. Series protocols are usually series-positive or series-negative. Series-positive means that a positive...
test result indicates the next test in sequence; a negative result makes the protocol negative. Series-negative means that a negative test result indicates the next test in sequence; a positive result makes the protocol positive.

To illustrate the techniques, we will first evaluate a simple test protocol consisting of two tests in parallel. We can model this protocol using established techniques (Turner, Frazer, Shepard, 1984; Turner, 1988; Turner et al., 1999). Measures of performance will be hit rate (HT), the probability of a positive result when the condition or disease is present and false alarm rate (FA), the probability of a positive result when the condition or disease is not present.

With two tests in parallel, there are four possible results (Fig. 2a). These are positive-positive, positive-negative, negative-positive, and negative-negative. We must establish a criterion, usually loose or strict, to determine if the protocol outcome is positive or negative. Regardless of criterion, all individuals would be defined as positive or negative.

**EXAMPLE**

In this example, we are testing 1000 children with a prevalence of hearing loss of 10%; therefore, 100 of the 1000 children have hearing loss. To simplify this example, we will make two assumptions. First, the tests are uncorrelated; that is, the tests have zero correlation. Test correlation is the tendency of two test to identify the same individuals as positive or negative. Zero correlation means that no such tendency exists. Second, all tests have the same performance, HT/FA = 80%/15%. The results of testing are shown in Figure 1b. Of the 100 children with hearing loss, 64 test positive on both tests, 32 have one positive and one negative result, and 4 test negative on both tests. With a strict criterion (both tests must be positive), the hit rate would be 64% (64/100). With a loose criterion (only one positive result is required), HT = 96% (96/100). Results for the children without hearing loss are also shown in Figure 1b. With a strict criterion, False Alarm Rate is 2% (20/900). With a loose criterion, FA = 28% (250/900). These results are summarized in Table 1.

As expected, a strict criterion reduces test protocol FA below that of the individual tests, which is good, but also reduces HT below that of the individual tests, which is bad. A loose criterion increases test protocol HT and FA above that of the individual tests. Thus, criterion can be used to manipulate HT and FA, but there is generally a trade-off: increasing HT increases FA and decreasing FA decreases HT. This trade-off is one reason why a test protocol may not be better than a single test. Which is better, the performance of the individual test (80%/15%), the performance of the protocol with a strict criterion (64%/2%), the performance of the protocol with a loose criterion (96%/28%)? There is no unequivocal answer to this question. The “best” testing strategy would depend on a number of factors including the particular testing objective, e.g., screening vs. diagnostic.

| **Table 1** Performance of a two-test protocol with a strict, loose, and cross-check criterion. |
|-----------------|--------|---------|--------|--------|---------|--------|
| **Criterion**   | HT(%)  | FA(%)   | UD(%)  | HT(%)  | FA(%)   | UD(%)  |
| Strict (+++)    | 64     | 2       | 0      | 72     | 9       | 0      |
| Loose (+)       | 96     | 28      | 0      | 88     | 21      | 0      |
| Cross-Check (++/-) | 94     | 28      | 26     | 66     | 13      | 13     |

Tests are combined in parallel and test performance is the same for each test (HT/FA = 80%/15%). The protocol is evaluated for zero (uncorrelated) and mid-positive test correlation. HT: Hit Rate; FA: False Alarm Rate; UD: Percent of tested individuals for whom no decision is made.
THE CROSS-CHECK PROTOCOL

Dissecting the Protocol

Jerger and Hayes proposed a single protocol using three tests (Fig. 1). This protocol is not a typical parallel or series protocol. With their protocol, all individuals receive the first two tests, behavioral and immittance. This is consistent with a parallel protocol, but inconsistent with a series protocol. The third test, BSER, is administered only when the first two tests disagree. Thus, not all individuals receive the third test, inconsistent with a parallel protocol.

With the cross-check protocol, all individuals receive the first two tests. As stated above, this is equivalent to two tests in parallel. What is unique about the cross-check protocol is the criterion used to determine if the protocol is positive or negative based upon these two test results. With a typical test protocol, a specific criterion is established that determines when the outcome of the protocol is considered positive or negative for the condition of interest, e.g., hearing loss. All tested individuals can be placed into one of two populations for whom the outcome is either positive or negative. The cross-check protocol does not do this. Again, quoting from Jerger and Hayes, “The basic operation of this principle is that no result be accepted until confirmed by an independent measure.” The cross-check protocol uses a criterion that creates three populations: one defined as positive because both tests were positive, one defined as negative because both tests were negative, and one population for which no decision is made because the two test results disagree. We will evaluate this unusual strategy by creating a two-test parallel protocol and using the same criterion used by the cross-check protocol. We will see that this strategy has advantages if an immediate decision is not required for every individual.

A second unique feature of the cross-check protocol is the use of a third test when the first two disagree. With this strategy, the protocol is called positive if the third test is positive. Effectively, the cross-check protocol requires two positive tests for the protocol to be positive. It turns out that the performance of the cross-check protocol is equivalent to a three-test parallel protocol with an intermediate criterion, i.e., two out of three tests must be positive. The proof of this is beyond the scope of this article. We will use a three-test parallel protocol to evaluate the performance of the cross-check protocol.

Performance of the Two-Test Protocol

We are interested in evaluating the two unique strategies contained within the cross-check protocol, as represented by our two-test and three-test protocols. We are not evaluating the performance of the actual protocol used by Jerger and Hayes in the 1970s. Thus, various test performances and test correlations will be used in the examples below without regard to any particular test.

In the cross-check protocol, all individuals receive the first two tests. This is represented by the two-test parallel protocol illustrated in Figure 2a. The criterion is the same as used in the cross-check protocol. Only those who test positive on both tests are considered positive; only those who test negative on both tests are considered negative. In this example (Fig. 2b), one thousand children are tested with a 10% prevalence of hearing loss. The tests are assumed to be uncorrelated and have the same performance, HT/FA = 85%/15%.

A total of 262 children have inconsistent test results and are considered neither positive or negative for hearing loss. Thus, no decision is made for 26% of those tested (Table 1). We can calculate a HT and FA for the remaining 74% for whom a decision has been made. For those with hearing loss, 64 tested positive and 4 negative, yielding a HT = 94%. For those with normal hearing, 20 tested positive and 650 negative, yielding a FA = 3%. Thus, implementing this feature of the cross-check protocol produced a test protocol performance of HT/FA = 94%/3%, which has a higher HT and lower FA than the individual tests (HT/FA = 80%/15%). Test protocol performance is clearly superior to that of the individual tests.

If a loose or strict criterion had been used with this two-test protocol, the undiagnosed population would have been eliminated; however, there would be the traditional trade-off between hit rate and false alarm rate. With a loose criterion, protocol HT would increase to 96%, but FA
would also increase to 39%. Likewise with a strict criterion, both HT and FA would be reduced relative the individual tests. In a sense, this unique strategy suggested by Jerger and Hayes provides the best of both worlds: the higher HT of a loose criterion and the lower FA of a strict criterion. Of course, we never get something for nothing. The cost of using this strategy is the failure to reach an immediate decision on every individual.

### Performance of the Three-Test Protocol

With the cross-check protocol, the use of the third test eliminates the undiagnosed population generated by the first two tests. How does this third test impact on the performance of the protocol? We can determine the performance of the cross-check protocol by evaluating a three-test parallel protocol using an intermediate criterion. The performance of a three-test protocol is shown in Table 2 for all possible criteria.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>HT(%)</th>
<th>FA(%)</th>
<th>UD(%)</th>
<th>HT(%)</th>
<th>FA(%)</th>
<th>UD(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strict (+++)</td>
<td>51</td>
<td>&lt;1</td>
<td>0</td>
<td>66</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Intermediate-CC (+++)</td>
<td>90</td>
<td>6</td>
<td>0</td>
<td>85</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Loose (+)</td>
<td>99</td>
<td>39</td>
<td>0</td>
<td>90</td>
<td>27</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: Performance of a three-test protocol with a strict, intermediate, and loose criterion.

Tests are combined in parallel and test performance is the same for each test (HT/FA = 80%/15%). The protocol is evaluated for zero and mid-positive test correlation. The intermediate criterion has the same performance as the three-test cross-check strategy. HT: Hit Rate; FA: False Alarm Rate; UD: Percent of tested individuals for whom no decision is made.

Particularly interesting is the result for the intermediate criterion, which is also the performance of the cross-check protocol. Note that the HT is greater than that of the individual tests (90% vs. 80%) and that the FA is lower than that of the individual tests (6% vs. 15%). Unlike a loose and strict criterion which move HT and FA in the same direction, the intermediate criterion increases HT and decreases FA. Clearly, this performance is better than that of the individual tests. Also note that the performance of the three-test cross-check protocol is not quite as good as the two-test protocol (90%/6% vs. 94%/3%). However, using three tests produces a decision for all individuals, unlike the two-test strategy that results in an undiagnosed population. Thus, the cost of an immediate decision is some possible reduction in performance.

### Test Correlation

In the examples thus far, tests were assumed to have zero correlation. Increasing test correlation will reduce the ability of a test protocol to manipulate HT and FA and, thus, the advantage of multiple tests over a single test. Limited data for audiological tests designed to distinguish cochlear from retrocochlear site-of-lesion indicate a mid-positive correlation may be an appropriate estimate (Turner, 1988). Mid-positive correlation is half way between zero correlation and maximum possible correlation. The advantages of the cross-check strategies may be reduced, or eliminated, with increasing test correlation.

We can recalculate the performance of the example test protocols using mid-positive
correlation. For the two-test protocol (Table 2), a strict criterion reduced HT and FA, but not as great as when tests have zero correlation (72%/9% vs. 64%/2%). For a loose criterion, HT and FA are increased, but, again, not as much as when tests are uncorrelated (88%/21% vs. 96%/28%). With the cross-check criterion, HT is increased (86% vs. 80% for individual test), but less than for uncorrelated tests (96%). Likewise, FA is reduced (10% vs. 15% for individual tests), but less than for uncorrelated tests (3%). Thus, increasing test correlation does reduce the advantage of the two-test protocol, but performance is still superior to the individual tests, even for a mid-positive correlation. The results are similar for the three-test protocol. The ability of the test protocol to modify HT and FA is reduced with mid-positive correlation. However, for an intermediate criterion, which corresponds to the cross-check strategy, HT is still greater and FA lower than that for the individual tests (85%/11% vs. 80%/15%). The superior performance of the two-test protocol relative the three-test protocol diminishes with increasing test correlation and is almost eliminated with mid-positive correlation (86%/10% vs. 85%/11%).

Varying Test Performance

In the examples thus far, all individual tests had the same performance (80%/15%). This is not what we would expect in an actual clinical situation. What is the impact of varying individual test performance on test protocol performance? The previous examples has been recalculated using individual test performance (HT/FA) of Test 1 = 70%/30%, Test 2 = 80%/15%, and Test 3 = 90%/10% (Table 3). These performance values are not intended to represent any particular tests, only to provide a range of values.

The two-test protocol still produced performance significantly better than either individual test for both zero and mid-positive correlation. As expected, there was no decision for a significant percentage of the test population (36% and 25%). A different result is obtained using three tests. Note that the performance using an intermediate criterion (90%/8%), which is equal to the cross-check performance, is only slightly better than that of the best individual test, Test 3 (90%/10%), for zero test correlation, and poorer (85%/12%) for mid-positive correlation. In this example, the most efficient test strategy would be to use just Test 3. There is little advantage to using additional tests to enhance performance; although, we could still manipulate HT and FA using a strict or loose criterion. We obtained this result because one test was significantly better than the other two. Thus, in a real clinical application, it should be recognized that the use of multiple tests may not significantly enhance performance if one test is much better than

| Table 3  Performance of a two-test protocol with a strict, loose, and cross-check criterion and a three-test protocol with a strict, intermediate, and loose criterion. |
|----------|--------|--------|--------|--------|--------|
| Criterion | HT(%)  | FA(%)  | UD(%)  | HT(%)  | FA(%)  | UD(%)  |
| Two Test  |        |        |        |        |        |        |
| T1=70/30  |        |        |        |        |        |        |
| T2=80/15  |        |        |        |        |        |        |
| Strict (+++) | 56   | 5    | 0     | 63    | 10    | 0    |
| Loose (+)  | 94    | 41   | 0     | 87    | 35    | 0    |
| Cross-Check (++/—) | 90  | 7    | 36    | 83    | 13    | 25   |
| Three Test |        |        |        |        |        |        |
| T1=70/30  |        |        |        |        |        |        |
| T2=80/15  |        |        |        |        |        |        |
| T3=90/10  |        |        |        |        |        |        |
| Strict (+++) | 50 | <1       | 0     | 60    | 5     | 0    |
| Intermediate-CC (+) | 90 | 8    | 0     | 85    | 12    | 0    |
| Loose (+)  | 99    | 46   | 0     | 95    | 38    | 0    |

Tests are combined in parallel and test performance is different for each test. For Test 1, HT/FA = 70%/30%; For Test 2, HT/FA = 80%/15%; For Test 3, HT/FA = 90%/10%. The protocol is evaluated for zero and mid-positive test correlation. The intermediate criterion has the same performance as the three-test cross-check strategy. HT: Hit Rate; FA: False Alarm Rate; UD: Percent of tested individuals for whom no decision is made.
the other tests. Also note that the performance of the three-test protocol is slightly poorer than that of the two-test protocol (90%/8% vs. 90%/7%) for zero correlation; however, it is actually slightly better (85%/12% vs. 83%/13%) for mid-positive correlation.

**DISCUSSION**

The cross-check protocol recommended by Jerger and Hayes utilizes two interesting strategies typically not found in more traditional test protocols. The first strategy requires that all initial tests agree before a decision is made. This produces a population for whom no decision is made. With a typical test protocol, a criterion is established that identifies all tested individuals as positive or negative. The advantage of this unique strategy is that it has the potential to significantly improve testing accuracy in the “diagnosed” population. Stated quantitatively, this strategy can produce a protocol hit rate that is significantly higher than that of the individual tests, and a protocol false alarm rate that is significantly lower than that of the individual tests. The disadvantage is that there will be an “undiagnosed” population.

This strategy of requiring test agreement is particularly useful if a decision can be delayed for some in the test population. In some situations, delaying a decision to permit additional testing may be reasonable. With children, additional time may permit maturation of the auditory system or the resolution of a transient condition, resulting in consistent test results. Of course, audiologists recognize the importance of early identification of hearing loss; however, a correct decision is more important than a quick decision. In fact, many audiologists use some form of this strategy when evaluating a difficult-to-test population. If all test results are consistent, the audiologist makes a decision and moves ahead with confidence. If test results are inconsistent, the individual is brought back for additional testing. The alternative would be to establish a specific criterion that would force a decision even with test disagreement. We have shown that making an immediate decision may result in poorer testing performance than delaying a decision for some in the test population. This is particularly true if retesting produces more consistent test results.

In some situations, retesting may not improve test agreement, or an immediate decision is necessary. Jerger and Hayes addressed this problem by referring their undiagnosed population for a third test, essentially a tie-breaker. Their total cross-check protocol reflects the second unique strategy. Their protocol is not a traditional parallel or series protocol. Its performance can be modeled as a three-test parallel protocol, but with an intermediate criterion, not the usual strict or loose criterion. In reality, the design of the cross-check protocol was determined by some very practical issues in the 1970s. The result was, however, a protocol that utilized an intermediate criterion, two out of three test must be positive. We have seen in this paper that an intermediate criterion has the potential to simultaneously increase hit rate and decrease false alarm rate, something a strict or loose criterion cannot do.

The two strategies extracted from the cross-check protocol have one important characteristic. They can produce test performance better than the individual tests being used. Sometimes this improvement in performance can be significant; however, the amount of improvement is diminished by greater test correlation. The primary application of the cross-check principle has been in the evaluation of pediatric and multiply-impaired populations. The data indicating mid-positive correlation are for tests that have little relevance to this application. Gans and Gans (1993) evaluated BOA, VRA, ABR, and the noise-tone-difference test (NTD) with multiply-impaired individuals. The NTD is an acoustic reflex based procedure for predicting threshold. Using statistical analysis, they found small correlations between these tests. While additional data are needed on test correlation, these results do suggest that the correlation for tests typically employed in the pediatric application may be significantly less than mid-positive. This suggests that the two strategies evaluated in this paper may be particularly useful with the pediatric population. It should also be noted that if the three tests vary significantly in performance, then the performance obtained with the two strategies may not be better than the best individual test.
The two strategies that were extracted from the cross-check protocol have a unique property. Both have the capability of simultaneously increasing hit rate and decreasing false alarm rate. The first strategy, only making a decision when all tests agree, is extremely powerful, but results in an undiagnosed population. If retesting this population in the future will result in greater test agreement, then this is a useful strategy, if a decision can be delayed.

The second strategy is the use of an intermediate criterion. It is more common for test protocols to employ a loose or strict criterion. The impact of these criterion on protocol performance is well understood. While they can be used to manipulate hit rate or false alarm rate, there is always the trade-off between hit rate and false alarm rate. They cannot simultaneously increase hit rate and reduce false alarm rate producing protocol performance clearly better than any individual test. Apparently an intermediate criterion has this capability, at least for the specific application in this paper. Even though more information is needed on the impact of intermediate criterion on protocol performance, audiologists should consider an intermediate criterion when using multiple tests.

Acknowledgment. The author would like to thank Dr. Barbara Cone-Wesson for her review of, and helpful comments about, this paper.

REFERENCES


To the Editor:

I have strong reactions to some of the points made in the editorial published in *JAAA*, Vol. 14, no 1, 2003. I agree with much of what you have to say about the connection between the laboratory and clinical practice. I would make the same case for the value of integrating differential thresholds for sensitivity and frequency, temporal integration, and the use of other psychoacoustic principles in the APD battery. Where we disagree is in the language you use to describe the Wilson et al. article. Specifically, the editorial states that that the Wilson article “reports normative data for a CD version of the MLD.” In fact the article reports central tendency data for the subjects in the study. The subjects are young adults, and only 24 of them.

It would have been more accurate for the editorial to state that the Wilson article provides findings on young adults who are considered normal. There are not nearly enough subjects, nor are the subjects from a widely distributed geographic, SES, educational level, etc. to be considered normative data. Further, the data are too limited to provide real clinical insights into the test-retest variability in a clinic population. The data are for young adults, there are no data for subjects older than 30 and none for subjects younger than 20. This is especially important because you link the MLD to children in your additional remarks, but there are no data related to children.

Additionally, the editorial states “MLD is reduced in children with APD, in some children with histories of COM.” The editorial needs to be accurate in stating that the MLD is reduced in some children with APD, and in some children with histories of COM. We do not actually know the prevalence of abnormal MLD data in any of these groups. The reader who takes for granted that what you write is truth might not see these subtle differences in language.

The Wilson article is a fine piece of research and important to the literature. However I think that the editorial is misleading.

Robert W. Keith
University of Cincinnati