

Three brief assessments of math achievement

Eric T. Steiner · Mark H. Ashcraft

Published online: 11 January 2012
© Psychonomic Society, Inc. 2012

Abstract Because of wide disparities in college students' math knowledge—that is, their math achievement—studies of cognitive processing in math tasks also need to assess their individual level of math achievement. For many research settings, however, using existing math achievement tests is either too costly or too time consuming. To solve this dilemma, we present three brief tests of math achievement here, two drawn from the Wide Range Achievement Test and one composed of noncopyrighted items. All three correlated substantially with the full achievement test and with math anxiety, our original focus, and all show acceptable to excellent reliability. When lengthy testing is not feasible, one of these brief tests can be substituted.

Keywords Math achievement · Brief assessment · Math anxiety · Wide Range Achievement Test · Math aptitude

Language-based research in cognitive science proceeds under the reasonable assumption that testing only those who are fluent (native) speakers of the target language will yield a sample that is both proficient and homogeneous in its knowledge of relevant information. As the logic goes, if college students are fluent speakers, they are competent enough—and so are comparable enough, for sampling purposes—to be able to judge which sentences are grammatical, to decide which letter strings are true words, to read and remember passages of text, and so forth.

It is increasingly clear that researchers in the area of numerical cognition cannot safely make the same assumption about the math knowledge of their college samples. Entering college students present with wide discrepancies

in their mathematics preparation, and many need remedial coursework before attempting college-level math; for example, recent evidence shows that only 23% of U.S. 12th graders are proficient in mathematics (National Mathematics Advisory Panel, 2008). For researchers, this means that a sample of college participants cannot be assumed to be homogenous with respect to knowledge and fluency in math, despite the sample's relative uniformity for language-based tasks. Such heterogeneity can undermine statistical outcomes and conclusions in a variety of ways, of course; for example, overly large within-group variability, because systematic variance due to differences in math knowledge or achievement has not been removed, can inflate the Type II error rate.

Given this circumstance, researchers in mathematical cognition should routinely determine individuals' level of math achievement or skill as a component of the research project, much the way math cognition researchers determine working memory capacity or math anxiety levels (e.g., Beilock, 2008; Beilock & DeCaro, 2007; Maloney, Risko, Ansari, & Fugelsang, 2010). Depending on circumstances, this information can be used to form separate groups (e.g., high vs. low achievement), to remove or partition variance using covariance or regression procedures, or to draw conclusions using the achievement information as corollary evidence in interpreting results (for examples, see Ashcraft & Kirk, 2001; LeFevre, DeStefano, Penner-Wilger, & Daley, 2006; Thevenot, Castel, Fanget, & Fayol, 2010). Any solution depends, however, on obtaining an estimate of the individuals' math achievement by administering some test. The question then becomes which test is suitable.

It is possible to use a full-fledged assessment—for instance, a standardized test—to determine math achievement. Although such approaches are thorough and gain from standardization procedures and norms, the tests can be quite costly and time consuming, and they often require special training to be administered. As an alternative, some researchers use a

E. T. Steiner · M. H. Ashcraft
University of Nevada Las Vegas,
Las Vegas, NV, USA

M. H. Ashcraft (✉)
Department of Psychology, University of Nevada,
Las Vegas, NV 89154, USA
e-mail: Mark.Ashcraft@unlv.edu

very brief test like the French kit (French, Ekstrom, & Price, 1963)—in particular, the two subtests that yield a four-page fluency test consisting of one- and two-column addition, subtraction, and multiplication problems. A disadvantage here is that the test assesses fluency, rather than math achievement, and consists only of three simple arithmetic operations, rather than the more thorough range of arithmetic and mathematics taught in K–12 curricula.

In our own work (e.g., Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007), we have often used the mathematics section of the *Wide Range Achievement Test: Third Edition* (WRAT3; Jastak Associates, 1993), a commonly used standardized test of math achievement. The test consists of 40 items of graded difficulty, to be administered in a 15-min session; the items include simple whole number arithmetic, problems with fractions and decimals, long division, percentages, and algebra with exponents and two unknowns. The test yields a composite score and is accompanied by age-appropriate norms that yield an interpretation about grade-appropriate attainment of mathematical achievement (Wilkinson, 1993).

In our work, we have scored college students' tests on a line-by-line basis, given that each line of the test presents progressively more difficult math problems to be solved, and then have examined performance as a function of the participants' levels of math anxiety. As is shown in Ashcraft and Krause (2007), performance on the first three lines of the WRAT3 is essentially perfect for college students, regardless of their math anxiety; these lines test whole-number arithmetic. Beyond that point in the test, however, performance starts to dip below the ceiling and starts to separate as a function of math anxiety. By the last line of the test, high math anxious participants average just under 1 item correct (out of 5), medium-anxious participants average 1.5 correct, and low-anxious participants average around 2.3 items correct. Overall, college-level math aptitude/achievement, as determined by grades or achievement scores, has a negative correlation with math anxiety ($r = -.31$; see Hembree, 1990, for this meta-analysis). In our present sample (Study 1) of 153 participants who completed both the WRAT3 and the shortened Mathematics Anxiety Rating Scale (SMARS; Alexander & Martray, 1989), the correlation was $r = -.35$ ($p < .001$).

We had two difficulties with the WRAT3 math test in its current form, which may mirror difficulties that others have (or will have) when they use it for other purposes. First, because the test is intended for use across all school ages, it has many items that are far too easy for college students (essentially the first three lines of the test; e.g., $1 + 1 = ?$; $3 \times 4 = ?$) and, thus, are a waste of time. Second, the test consumes 15 min, which, in a typical experimental session of 1 h, is a considerable time investment. It would be a substantial improvement, we decided, if a shorter assessment of math achievement could be developed. Thus, our interest was in developing a shorter test of math achievement that

maintained a good predictive relationship with math anxiety for our research and, of course, was still a reliable and valid substitute for the full WRAT3. Such a test could be generally useful to all researchers in the area, we reasoned, as a way of determining participants' level of math achievement.

Study 1

Method

Participants A total of 180 individuals participated in this study. All participants (both studies) were recruited from the departmental participant pool and received course credit for participation. Ages ranged from 18 to 53 years ($M = 22.40$, $SD = 5.79$). There were 105 females and 75 males. Ethnic breakdown was as follows: 104 Caucasian, 29 African-American, 26 Asian/Polynesian, 16 Hispanic, and 5 other. All were fluent in English.

Procedure Of the 180 participants, 15 failed to complete the entire WRAT3 assessment (yielding $n = 165$ for those analyses), and 12 more failed to complete both the WRAT3 and the SMARS (yielding $n = 153$ for those analyses). We refer to the 10 items extracted from the WRAT3 as the Brief Mathematics Assessment-1 (BMA-1) and describe next how these items were selected for use in the brief assessment.

The 10 extracted items for the BMA-1 represented problems of increasing difficulty, as is the case in the original WRAT3. The first four items tested the four basic arithmetic operations, using whole numbers (single- and two-digit values). These items were included not so much to provide predictive power for math achievement or math anxiety as to mimic the overall structure of the WRAT3 and to motivate participants to try and complete as many items as possible; that is, a questionnaire that contains only difficult items may discourage some individuals and lead them to quit prematurely.

The remaining six items for the BMA-1 assessment were chosen to maximize differences in performance as a function of math anxiety, as follows. We categorized our participants into low, medium, and high math anxiety groups on the basis of their SMARS scores, using our standard group membership criteria (e.g., Ashcraft & Krause, 2007).¹ As such, 34 were low math anxious, 54 were medium math

¹ In our work with the SMARS assessment, the overall mean is approximately 36.0, with a standard deviation of 16. We use these values for determining math anxiety group membership. Specifically, for membership in the low math anxiety group, a participant must score at least one 1 *SD* below the mean—in other words, 20 or lower. For the high math anxiety group, the score must be at least 1 *SD* above the mean—52 or higher. For the medium-anxiety group, a score must fall in the range of 1 *SD* centered on the mean—in other words, from 28 to 42 (36 ± 0.5 *SD*).

anxious, and 37 were high math anxious. We then examined the percentage of participants from each anxiety group who completed each of the 40 WRAT3 items correctly. We then selected the final 6 items for the BMA-1 so that each item demonstrated the greatest range between low and high math anxious groups in terms of percent correct. As an example, we chose WRAT3 item 21 ($2\frac{1}{2} + 1\frac{1}{2} = ?$) because it was answered correctly by 57% of the high math anxious individuals, 80% of the medium math anxious individuals, and 94% of the low math anxious individuals. Conversely, an item was not selected if it showed little variability in performance across the anxiety groups (whether high or low accuracy), since such items do not provide useful information about differences in math achievement or math anxiety.

The 10 BMA-1 items selected were WRAT3 items 7, 9, 10, 18, 21, 27, 32, 35, 37, 38, and 39. (Note that items 37 and 38 correspond to our 9th item; on the WRAT3, these are the two scored answers to a two-part question, an algebra problem with two unknowns.) Because of copyright restrictions, the test items cannot be provided here. Those with access to the test, however, may simply duplicate the items from the WRAT3 and use them as a brief assessment.

Results

Statistical analysis Correlations were computed on individuals' scores on the WRAT3, BMA-1, and SMARS. An item analysis was also calculated for the internal consistency of the BMA-1 items.

The correlation between the overall WRAT3 and the 10 extracted items on our BMA-1 brief test was quite substantial, $r = .90$ ($n = 165$; all correlations presented here attained the .001 significance level). On the basis of this sample, Cronbach's alpha was .78 for the BMA-1, above the traditionally accepted level of .70 (Nunnally & Bernstein, 1994). Note that the item analysis involved 11 items, since the 9th item of the BMA-1 is scored for two answers, one for each of the two unknowns. For the 153 participants who completed both the WRAT3 and the SMARS, there was an inverse correlation between the two measures, $r = -.35$, corresponding very favorably with the meta-analysis result between math anxiety and math achievement ($r = -.31$) reported in Hembree (1990). For these same participants, the correlation between the 10 items on our BMA-1 and the SMARS was very similar, $r = -.34$.

For those who might wish to convert an individual's BMA-1 score to the original scale of the WRAT3, the regression equation in our analysis was the following:

$$\begin{aligned} \text{Predicted WRAT3 score} &= 2.233(\text{BMA-1 score}) \\ &+ 14.373, \text{ standard error of the estimate} = 2.69, R^2 = .796. \end{aligned}$$

Study 2

An additional difficulty with WRAT3 is that it has now been superseded by a new edition of the test, the WRAT4 (Wilkinson & Robertson, 2006). This suggests that the WRAT3 and, therefore, the items for the BMA-1 will become less accessible. As such, we decided to generate a parallel short assessment on the basis of the new WRAT4 instrument, to be called the BMA-2. Because items from the WRAT4 cannot be furnished here either, again due to copyright restrictions, we also decided to write a third brief assessment, one that would not be subject to copyright restrictions. For reasons noted below, we based this third assessment, to be called the BMA-3, on the structure of problems found in the WRAT3.

Comparison of the older WRAT3 and the new WRAT4 suggests that the newer test has been simplified. While 26 of the WRAT4 problems are identical to those on the WRAT3, another 6 have been revised. One of the revised problems is now harder, but the other 5 have been simplified (e.g., in a long multiplication problem, the two-digit multiplier was changed from 96 to 45). Eight new problems appear on the WRAT4, mostly in the early lines where lower difficulty problems are positioned. Significantly, of the 5 most difficult problems on the WRAT3 (last line of the test), only 1 remains on the WRAT4; the problems that were dropped from the last line were the algebra problem with two unknowns, the factoring problem with exponents, and the function problem. Because of these simplifications, we expected the WRAT4 to be somewhat less challenging, especially in view of our earlier results on math anxiety. As a consequence, we modeled the structure and the problems for our own BMA-3 on the more challenging WRAT3, to preserve some of the more difficult computations found on that earlier test.

Method

Participants A total of 189 individuals were tested in this study. Ages ranged from 17 to 49 years ($M = 21.37$, $SD = 4.75$). There were 104 females, 80 males, and 5 undisclosed. Ethnic breakdown was as follows: 98 Caucasian, 36 Asian/Polynesian, 23 African-American, 25 Hispanic, and 6 unspecified.

Procedure A total of 184 participants completed both the WRAT4 and the SMARS, and all but 7 completed the Brief Math Assessment-3 (BMA-3) that we developed. Note that, unlike the procedure followed in Study 1, participants here completed both the WRAT4 and the 10-item BMA-3. In other words, the BMA-3 scores were not extracted from performance on the WRAT4, since the items on the BMA-3 are different from those on the WRAT4. The

scores on the BMA-2, however, were extracted from the overall WRAT4 tests, following the procedure used for the BMA-1 in Study 1.

To compensate for the perceived simplification of the WRAT4, we decided to use only 2 easy, or “motivational,” items on the BMA-2 (instead of 4, as in the BMA-1). For the remaining items, we followed the procedures used in Study 1, examining performance across all 40 items on the WRAT4 separately for the three math anxiety groups and selecting items that showed the greatest range between low and high math anxious groups in terms of percent correct. The final 10 items that we extracted from the WRAT4 for use on the BMA-2 were items 8, 19, 30, 31, 33, 34, 35, 37, 38, and 40. Note that this selection of items concentrates more heavily on the last two lines of the test (items 31 and higher), where the more difficult items are positioned, in agreement with our judgment that the WRAT4 is somewhat simplified over the earlier version.

For the BMA-3, we wrote 10 items, distributed over three rows of increasing difficulty. The items were developed to be similar in difficulty and type of math problems as seen in the WRAT3 and BMA-1 but used different numbers (e.g., the BMA-1 item $2\frac{1}{2} + 1\frac{1}{2}$ became $3\frac{1}{2} + 2\frac{1}{2}$). The BMA-3’s first row of four questions has very easy computations of subtraction, addition, multiplication, and division. As with the BMA-1 items, these were not intended to provide predictive power of math achievement or math anxiety but were included to motivate participants to try and complete as many items as possible. The remaining items also resemble the structure and increasing difficulty of the later items in the WRAT3 and BMA-1, including the algebra problem with two unknowns and the factoring problem with exponents. The Appendix presents the BMA-3 and is free for readers to use or revise as they see fit.

We told participants they had 10 min to complete the BMA-3 and that they should attempt all items if possible. Generally, participants finished the test in 6–8 min; they seldom answered more problems after that.

Results

Statistical analysis Correlation analyses were performed on individuals’ scores on the WRAT4, SMARS, BMA-2, and BMA-3. An item analysis was also calculated for the internal consistency of the BMA-2 and -3 items.

The BMA-2 test correlated significantly with the WRAT4, $r = .90$, on the basis of the 184 available scores. Cronbach’s alpha was .78 for this 10-item test (for comparison purposes, Cronbach’s alpha was .84 on the WRAT4 for our sample). The typical inverse relationship between math anxiety and math achievement scores was obtained both on the WRAT4 ($r = -.31$, $n = 184$) and on the BMA-2 ($r = -.33$, $n = 184$). Table 1 presents the mean percentage correct (and standard deviations) of the 10 items on the BMA-2, along with the corresponding item number from the WRAT4, both for our overall sample and also separately for the three math anxiety groups.

For the BMA-3 test we constructed, there was an overall correlation of .66 with the WRAT4 ($n = 177$). On the basis of 182 participants who completed the BMA-3, Cronbach’s alpha was .69 when scores on the 11-point scale were analyzed (given that item 9 has two scorable answers). And the BMA-3 scores were negatively correlated with math anxiety scores ($r = -.27$, $n = 177$). Although this correlation is somewhat lower than the BMA-1 correlation with math anxiety from Study 1 ($r = -.34$), it does not differ statistically from that value ($z = 0.75$, $p = .23$). Table 2 presents values for the BMA-3 equivalent to those in Table 1.

Table 1 Means and standard deviations for items on the BMA-2 for the full sample and separately for low, medium, and high math anxiety groups

Item	WRAT4 #	Full Sample		Low Anxiety		Medium Anxiety		High Anxiety	
		M	SD	M	SD	M	SD	M	SD
		(n=184)		(n=40)		(n=44)		(n=41)	
1	8	.96	.20	.97	.16	.93	.25	.93	.26
2	19	.87	.33	.95	.22	.91	.29	.78	.42
3	30	.63	.48	.82	.38	.68	.47	.51	.51
4	31	.71	.46	.85	.36	.73	.45	.56	.50
5	33	.64	.48	.82	.38	.59	.50	.49	.51
6	34	.62	.47	.80	.40	.66	.48	.51	.51
7	35	.55	.50	.68	.47	.68	.47	.44	.50
8	37	.61	.49	.80	.40	.66	.48	.54	.50
9	38	.37	.48	.50	.51	.41	.50	.32	.47
10	40	.40	.49	.60	.50	.39	.49	.37	.49
Mean		6.37	2.60	7.80	2.06	6.64	2.41	5.44	2.90

Table 2 Means and standard deviations for items on the BMA-3 for the full sample and separately for low, medium, and high math anxiety groups

Item	Full Sample (n=182)		Low Anxiety (n=37)		Medium Anxiety (n=44)		High Anxiety (n=40)	
	M	SD	M	SD	M	SD	M	SD
1	.98	.15	1.0	.00	.95	.21	.97	.16
2	.88	.32	.92	.28	.91	.29	.87	.33
3	.98	.15	1.0	.00	.98	.15	.95	.22
4	.97	.16	1.0	.00	.95	.21	1.0	.00
5	.86	.34	.95	.23	.86	.35	.75	.44
6	.75	.44	.95	.23	.77	.42	.60	.50
7	.45	.50	.62	.49	.43	.50	.42	.50
8	.37	.48	.49	.51	.41	.50	.30	.46
9a*	.61	.49	.78	.42	.59	.50	.52	.50
9b*	.61	.49	.78	.42	.59	.50	.52	.50
10	.15	.36	.35	.48	.11	.32	.12	.33
Mean	7.61	2.06	8.84	1.80	7.57	2.11	7.05	2.17

Items 9a and 9b refer to the two answers to question 9 on BMA-3. Interestingly, all participants who got the correct answer to 9a also answered correctly on 9b; hence, all means and *SDs* for 9a and 9b are identical. Nonetheless, the mean and *SD* on the test are scored on 11 points overall

Incorrect answers were approximately three times as common as omissions on the BMA-3, probably not surprising given that most participants finished within 8 min (this was also the general pattern of errors on the WRAT3 and WRAT4). Blank answers were common on the last two test items (and on the last line in the larger WRAT tests); some participants even wrote comments on these items, such as “I don’t remember how to do this.” The slight drop in accuracy on item 2 came from a perseveration tendency in which participants subtracted, just as they did (correctly) on item 1. Item 7, with mixed fractions and borrowing, was especially difficult, and the overwhelmingly popular wrong answer to item 8 was $\frac{1}{4}$.

For those who might wish to convert an individual’s BMA-2 or -3 score to the original scale of the WRAT4, the regression equations in our analyses were the following:

$$\text{Predicted WRAT4 score} = 1.758(\text{BMA} - 2 \text{ score})$$

$$+ 21.16, \text{ standard error of the estimate} = 2.23, R^2 = .808.$$

$$\text{Predicted WRAT4 score} = 1.584(\text{BMA} - 3 \text{ score})$$

$$+ 20.364, \text{ standard error of the estimate} = 3.76, R^2 = .435.$$

The correlation between BMA-3 and WRAT4 scores (.66) was significantly lower here ($z = 2.12, p = .017$) than the BMA-1 to WRAT3 correlation found in Study 1 (.90), which we attribute to the simplification of items on the WRAT4.² That is, the items for the BMA-3 were modeled on the types of math operations, and the difficulty, found on the WRAT3. Performance on the BMA-3, however, is being compared with performance on the simpler WRAT4, thus perhaps explaining the significant reduction in the correlation. In keeping with this explanation, we computed participants’ mean performance on line 8 of the WRAT4, as a comparison with the values presented above. On the WRAT3, our earlier research (Ashcraft & Krause, 2007) found mean correct (out of 5) on line 8 was 2.3, 1.5, and 0.9, respectively for low-, medium-, and high-anxious groups; in terms of percent correct, these were 46%, 30%, and 18%. In the present WRAT4 data, the comparable means were 2.8, 2.3, and 2.0, again respectively from low to high anxious (in percentages, 56%, 46%, and 40%). In other words, correct performance on the WRAT4’s most difficult line of problems was from 10 to 22 percentage points higher than on the same line of the WRAT3 (and, recall, with very large sample sizes).

Conclusion

The three BMA tests presented here are short and effective ways of assessing individuals’ math achievement. Each test requires no more than approximately 6–8 min for administration, demonstrates a strong to very strong relationship to math achievement as measured by the standardized WRAT3 or WRAT4 instrument, and shows suitable to excellent reliability. Each test covers the range of arithmetic and math assessed in the larger WRAT tests. Unlike the French kit (French et al., 1963), which tests only three simple arithmetic operations, these three brief tests sample items more broadly in arithmetic (whole numbers and fractions) and algebra computation procedures. In addition, the tests allow for the prediction of math anxiety. The tests are appropriate for use in research settings where a brief assessment of math achievement is needed or, possibly, in any setting where a brief screening on math achievement would be useful before deciding whether to administer a more lengthy standardized test. Users will find the items in the tests very similar in terms of the type of mathematical questions involved and, aside from the simplifications noted in the WRAT4, in their level of difficulty.

² Note also that the BMA-1 to WRAT3 correlation is high because responses on the BMA-1 were extracted from the WRAT3 and represent one quarter of the items on the WRAT3.

Author Note Eric T. Steiner and Mark H. Ashcraft, Department of Psychology, University of Nevada, Las Vegas.

Eric T. Steiner is now in the Social & Behavioral Science Division, Central Arizona College, Coolidge, AZ.

Appendix

Brief mathematics assessment-3

Participant # _____

<p>1.</p> $\begin{array}{r} 42 \\ -21 \\ \hline \end{array}$	<p>2.</p> $\begin{array}{r} 56 \\ +17 \\ \hline \end{array}$	<p>3.</p> $\begin{array}{r} 8 \\ \times 5 \\ \hline \end{array}$	<p>4.</p> $\frac{9}{3} = \underline{\hspace{2cm}}$
<p>5.</p> $3\frac{1}{2} + 2\frac{1}{2} = \underline{\hspace{2cm}}$	<p>6.</p> $\begin{array}{r} 4\frac{1}{4} \\ 3\frac{1}{8} \\ + 2\frac{1}{2} \\ \hline \end{array}$	<p>7.</p> $\begin{array}{r} 8\frac{1}{4} \\ -5\frac{2}{3} \\ \hline \end{array}$	
<p>8.</p> <p>Write as a common fraction in lowest terms:</p> <p>.025 = _____</p>	<p>9.</p> $5j - w = 18$ $4j - w = 14$ <p>$j = \underline{\hspace{1cm}}$ $w = \underline{\hspace{1cm}}$</p>	<p>10.</p> <p>Reduce:</p> $\frac{P^2 + P}{P^2} \cdot \frac{2P - 2}{P^2 - 1}$ <p>Answer: _____</p>	

References

- Alexander, L., & Martray, C. R. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development*, 22, 143–150.
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, 130, 224–237.
- Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, 14, 243–248.
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*, 17, 339–343.
- Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 983–998.
- French, J. W., Ekstron, R. B., & Price, I. A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21, 33–46.
- Jastak Associates. (1993). *Wide Range Achievement Test, Rev. 3*. Wilmington, DE: Wide Range Inc.
- LeFevre, J., DeStefano, D., Penner-Wilger, M., & Daley, K. E. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology*, 60, 209–220.
- Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition*, 114, 293–297.
- National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, D.C.: U.S. Department of Education. Available at www.ed.gov/pubs/edpubs.html
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Thevenot, C., Castel, C., Fanget, M., & Fayol, M. (2010). Mental subtraction in high- and lower skilled arithmetic problem solvers: Verbal report versus operand-recognition paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1242–1255.
- Wilkinson, G. S. (1993). *Manual for the Wide Range Achievement Test* (3rd ed.). Wilmington, DE: Wide Range.
- Wilkinson, G. S., & Robertson, G. J. (2006). *WRAT4 Wide Range Achievement Test professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.