

What Impact Does Calculator Use Have On Test Results?

Edward W. Wolfe

Calculators are commonly used in mathematics and science instruction. In fact, over 20 years ago, two studies commissioned by the College Board (Kupin & Whittington, 1988; Pfeiffenberger & Zolandz, 1989) indicated that, at that time, nearly all math and science college instructors permitted use of calculators for all types of course work and at least some tests. Those studies also indicated that, at that time, calculators were not commonly permitted on large-scale assessments. Since then, policies have changed, and calculator use is permitted on many large-scale assessments including the Stanford Achievement Test, the ACT, the SAT II, the NAEP, and PRAXIS, in addition to statewide assessments. Since that time, calculators have not only become commonplace, but their capacities have been dwarfed by the features offered by modern desktop computers and the mathematical programs that are available to students.

A common concern relating to calculator use on large-scale assessments is whether permitting their use somehow changes the

meaning of the scores that examinees receive. Examinee use of calculators could, potentially, completely change the nature of the construct being measured by the test items. For example, consider a test that is designed to measure basic arithmetic (e.g., addition, subtraction, multiplication, and division). If calculator use is not permitted on such a test, then examinees must rely on their arithmetic skills to answer the test items correctly. On the other hand, if calculator use is permitted, then examinees must rely on their skills at entering arithmetic equations into the calculator to answer the items correctly. Obviously, these are two very different skills, so the test scores would be interpreted quite differently, depending on whether calculators are permitted or not on such an assessment.

A common concern relating to calculator use on large-scale assessments is whether permitting their use somehow changes the meaning of the scores that examinees receive.

Calculator functionality may also be a relevant consideration. If the type of calculator is not standardized when calculator use is permitted, then it is possible that examinees who use different calculators and obtain the same test

score may demonstrate quite different levels of skill. For example, consider a test item that requires examinees to make predictions based on a formula. Examinees who use a basic calculator must be able to describe the trend based on results that they obtain by entering sets of numbers into the formula and extrapolating based on those results. On the other hand, examinees who use a graphing calculator can see an entire range of values for the equation in question by entering the formula into the calculator. Again, the skills being demonstrated by these two examinees are quite different.

Research about calculator use on tests measuring mathematics knowledge and skills that was conducted prior to the year 2000, suggested several effects of the calculator. Results were mixed concerning whether calculator use impacts total scores, with some researchers finding that calculator use increased test scores (Bridgeman, Harvey, & Braswell, 1995; Colton, 1997; Long, Reys, & Osterlind, 1989; Morgan & Stevens, 1991), and some studies finding no test score increases (Ansley, Spratt, & Forsyth, 1989; Dye, 1981; Forsyth & Ansley, 1982). However, there is at least some evidence that the impact of calculator use was different for different types of items. Some of those studies

indicated that calculator use improves performance on computational items (Dye, 1981; Lawrence & Dorans, 1994; Lloyd, 1991) while having less impact on problem solving items (Long, et al., 1989). In addition, it seems that graphing calculators were more likely to be associated with score increases (Morgan & Stevens, 1991) and that calculator use had little impact on the speed with which students completed the test (Lloyd, 1991).

Concerning differential impact on examinees, these early studies suggest that Latinos students, and minorities in general, may have benefited more from the use of calculators (Bridgeman, et al., 1995; Colton, 1997; Gao, 1997); that males might benefit more than females (Colton, 1997; Morgan & Stevens, 1991); and students with more calculator experience might receive more benefit (Bridgeman, et al., 1995). Research results were inconsistent concerning student ability, with a couple of studies suggesting a differential benefit favoring low ability students (Colton, 1997; Gao, 1997) and one study suggesting no differential effect for student ability (Bridgeman, et al., 1995).

Since the year 2000, research concerning calculator use has continued, and many of these

results have been replicated with more recent samples of students. For example, studies have continued to reveal mixed results concerning total score increases. While one study revealed a positive impact on total score for calculator users (Loveless, 2004) another study indicated that the magnitude of that advantage has decreased over time and are now at a level that is ignorable (Brooks, et al., 2003). Similarly, although some studies suggest that calculator use is more beneficial on some types of items, at least one study reveals no

Studies have continued to reveal mixed results concerning total score increases.

apparent differences. Specifically, several studies suggest that calculator use is favored and performance is improved on computation and graphing items while use and score impact is not as apparent on items measuring reasoning and conceptual understanding (Ellington, 2003; Scheuneman & Camera, 2002; Schwarz, Rich, Arenson, Podrabsky, & Cook, 2002). On the other hand, one study found no differential performance on items that were judged by experts to be susceptible to calculator use (Brooks, et al., 2003). Similarly, at least one of the recent studies indicates a differential benefit in favor of

Hispanic and black students when these students use calculators (Loveless, 2004), while another study suggests universal improvements for some gender, ethnicity, income, achievement, and course experience groups (Hanson, Brown, Levine, & Garcia, 2001).

Concerning the nature and type of calculator use, the recent studies have provided new insights into the impact of calculator use on mathematics tests. For example, one study reveals that judicious use of the calculator can improve test scores, while under or over use is associated with lower test performance. Scheuneman and Camara (2002) found that those who use calculators on one-third to one-half of the items on the SAT see the greatest benefit, in comparison with those who use calculators more or less than this. In addition, that study indicated that those using calculators most often were less likely to complete the test and, supported by the results of another study (Schwarz, et al., 2002), that use of graphing calculators was more predictive of higher test scores than use of other types of calculators. However, additional studies provided conflicting results. For example, one study suggests no impact on test scores associated with calculator type (Hanson, et al., 2001) while another study provides

no evidence of impact of calculator use on test completion rates (Schwarz, et al., 2002).

The more recent studies have also provided evidence that the way calculators are used outside of the test is probably as important as whether calculators are used during the testing session. For example, several studies have documented higher test scores for students who use calculators regularly during mathematics instruction (Chazan, et al., 2007; Heller, Curtis, Jaffe, & Verboncoeur, 2005; Scheuneman & Camera, 2002). In addition, one study revealed higher test scores for students of teachers who received training about how to use calculators when delivering mathematics content to students (Heller, et al., 2005).

Clearly, prior research indicates that calculator use may increase student scores on mathematics tests, particularly when the test contains a large proportion of computation items. In addition, calculator use may be more beneficial for historically disadvantaged students. However, it is also clear from the most recent studies that judicious use of the calculator, use of the right type of calculator, and integration of the calculator into mathematics instruction are keys to maximizing the positive impact of allowing

students to use calculators on mathematics tests.

References

- Ansley, T. N., Spratt, K. F., & Forsyth, R. A. (1989). The effects of using calculators to reduce the computational burden on a standardized test of mathematics problem solving. *Educational and Psychological Measurement*, 49, 277-286.
- Bridgeman, B., Harvey, A., & Braswell, J. (1995). Effects of calculator use on scores on a test of mathematical reasoning. *Journal of Educational Measurement*, 32, 323-340.
- Brooks, T., Case, B. J., Cerrillo, T., Severance, N., Wall, N., & Young, M. J. (2003). Calculator use on Stanford Series Mathematics Tests. (Assessment Report). San Antonio, TX: Pearson.
- Chazan, D., Leavy, A. M., Birky, G., Clark, K., Lueke, M., McCoy, W., et al. (2007). What NAEP can (and cannot) tell us about performance in algebra. In P. Kloosterman & J. F. Lester (Eds.), *Results and interpretations of the 2003 mathematics assessment of the National Assessment of Educational Progress*. Reston, VA:

- National Council of Teachers of Mathematics.
- Colton, D. A. (1997). *Monitoring calculator implementation for the ACT and PLAN*. Paper presented at the American Educational Research Association, Chicago, IL.
- Dye, D. L. (1981) The use and non-use of calculators on assessment testing. St. Paul, MN: Minnesota State Department of Education.
- Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal of Research in Mathematics Education*, 34, 433-463.
- Forsyth, R. A., & Ansley, T. N. (1982). The importance of computational skill for answering items in a mathematics problem-solving test: Implications for construct validity. *Educational and Psychological Measurement*, 42, 257-263.
- Gao, X. (1997). *Examining calculator effects on subgroup mathematics performance*. Paper presented at the American Educational Research Association, Chicago, IL.
- Hanson, K., Brown, B., Levine, R., & Garcia, T. (2001). Should standard calculators be provided in testing situations? An investigation of performance and preference differences. *Applied Measurement in Education*, 14, 59-72.
- Heller, J. I., Curtis, D. A., Jaffe, R., & Verboncoeur, C. J. (2005) Impact of handheld graphing calculator use on student achievement in algebra 1. Oakland, CA: Heller Research Associates.
- Kupin, J. L., & Whittington, B. R. (1988) A survey of the use of hand calculators and microcomputers in college mathematics classes. (Report No. 88-6). New York, NY: College Board.
- Lawrence, I., & Dorans, N. J. (1994) Optimal use of calculators on a mathematics test: Effect on item difficulty and score equating. (Report No. RR-94-40). Princeton, NJ: Educational Testing Service.
- Lloyd, B. H. (1991). Mathematics test performance: The effects of item type and calculator use. *Applied Measurement in Education*, 4, 11-22.
- Long, V. M., Reys, B., & Osterlind, S. J. (1989). Using calculators on achievement tests. *Mathematics Teacher*, 82, 318-325.
- Loveless, T. (2004). *Computational skills, calculators, and*

- achievement gaps: An analysis of NAEP items.* Paper presented at the American Educational Research Association, San Diego, CA.
- Morgan, R., & Stevens, J. (1991) Experimental study of the effects of calculator use on the Advanced Placement Calculus Examinations. (Report No. RR-91-5). Princeton, NJ: Educational Testing Service.
- Pfeiffenberger, G. W., & Zolandz, A. M. (1989) Surveys of the use of hand calculators and microcomputers in college preparatory and college science classes. (Report No. 89-4). New York, NY: College Board.
- Scheuneman, J., & Camera, W. J. (2002) Calculator Use and the SAT® I Math. (Report No. RN-16). New York, NY: College Board.
- Schwarz, R., Rich, C., Arenson, E., Podrabsky, T., & Cook, C. (2002). *An analysis of differential item functioning based on calculator type.* Paper presented at the National Council on Measurement in Education, New Orleans, LA.