
CONCEPTUAL ARTICLES

Calculator-Available Assessments: The Why, What, and How

David E. Meel

University of Pittsburgh

In the *Curriculum and Evaluation Standards for School Mathematics*, the National Council of Teachers of Mathematics (1989) advocated that "appropriate calculators should be available to all students at all times" (p. 8) and that calculators (or any instructional tool) should be allowed in the assessment of those students. In response, large-scale assessments such as the National Assessment of Educational Progress, Scholastic Achievement Test, American College Testing, Advanced Placement, and a variety of state-mandated tests moved to incorporate the allowance of calculator technologies in the assessment scheme. Although rhetoric and large-scale assessments have embraced the allowance of technology, classroom teachers may have difficulty developing their own assessments that permit calculator use. This article discusses the factors involved in the design and implementation of calculator-available assessments. In particular, four issues associated with calculator-available assessments are examined: (a) the benefits and problems associated with calculator use in assessment, (b) the components of various types of calculator-available assessments, (c) the construction of calculator-available assessments, and (d) the need for additional examination of items included on calculator-available assessments. From this examination, teachers and test developers are provided with a set of recommendations that focus on developing calculator-available assessments and minimizing difficulties arising at the item level.

In the past decade, large-scale standardized assessments of mathematical achievement have shifted to permitting the use of calculators. In particular, the National

Requests for reprints should be sent to David E. Meel, Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403. E-mail: meel@math.bgsu.edu

Assessment of Educational Progress (NAEP), Scholastic Achievement Test, and Advanced Placement exams, to name a few, as well as a variety of state-mandated tests, have begun the process of incorporating the calculator into their assessments (Braswell, 1992; Bridgeman, Harvey, & Braswell, 1993; Greenes & Rigol, 1992; Hopkins, 1992; Kouba & Swafford, 1989; Lienwand, 1992; Mullis, Dossey, Owen, & Phillips, 1991; Pandey, 1986; Payne, 1992). American College Testing decided to allow students to use calculators on the mathematics portion of the test beginning in the fall of 1996 ("ACT Allows Calculators," 1996). As a result, teachers need to know how to develop tests that match the current view of school mathematics, maintain equity between groups, reflect student achievement, and promote valid inferences.

One possible source of information relevant to designing classroom assessments has been the documents produced by the National Council of Teachers of Mathematics (NCTM), the world's largest mathematics education organization. The NCTM (1989) clearly supported inclusion of technological tools both in the classroom and during assessment when, in the *Curriculum and Evaluation Standards for School Mathematics*, it stated, "appropriate calculators should be available to all students at all times" (p. 8). Later, in the same document, the NCTM stated

When these materials [manipulatives, calculators, and computers] are used during instruction, they should be available during assessment, as long as their use is consistent with the purpose of the assessment. For example, if students routinely use calculators for solving problems in class, they should also be able to use calculators during assessments of their problem-solving abilities. (p. 195)

However, the *Curriculum and Evaluation Standards for School Mathematics* does not provide a set of recommendations to teachers for designing their own assessments that permit calculator use. The most recent NCTM (1995) document, *Assessment Standards for School Mathematics*, identified standards that help guide teachers in examining their assessment practices and in developing new assessment techniques. Unfortunately, the only mention of issues concerning the incorporation of technology into the assessment of students was reminiscent of those described in the *Curriculum and Evaluation Standards for School Mathematics*. Therefore, *Assessment Standards for School Mathematics*, although quite useful at many different levels, did not provide teachers or test designers with a clear-cut set of recommendations as to considerations necessary for the implementation of technology into the assessment. This leaves assessors with an impetus but no means or direction for implementation. This article attempts to fill this gap by outlining the benefits, problems, components, and additional considerations surrounding the implementation of calculator-available assessments. In addition, this article provides a short list of recommendations to help teachers and test developers in designing calculator-available assessments.

BENEFITS AND PROBLEMS ASSOCIATED WITH CALCULATOR USE IN ASSESSMENT

A number of benefits accrue from the inclusion of calculators in the classroom and in the domain of assessment. The inclusion of calculators as a tool in the assessment scenario ensures that students do not spend an inordinate amount of time on computation during testing (Hopkins, 1992). Calculators also open an opportunity to insert realistic problems into tests, thereby making them more reliable measures of students' conceptual understanding (Hopkins, 1992; Wilson & Kilpatrick, 1989). In addition, the implementation of the calculator allows the teacher to approach topics in new ways. When the focus of the lesson is on the underlying concepts, the calculator permits students to forgo longhand computations while maintaining a focus on exploring, comparing, and experiencing the mathematics (Reys, 1989).

Another positive impact of including calculators has been that students tend to develop better than average mathematical capabilities and understandings of certain domains, namely, problem solving, estimation, decimals, measurement, word problems, and operations (Finley, 1992; Hembree & Dessart, 1986, 1992; Kouba & Swafford, 1989; Loyd, 1991; Mullis et al., 1991; Reys, 1989; Silverman & Muehl, 1989; G. Wheatley & Shumway, 1992). These increased capabilities are not at the expense of computational ability or basic skills as long as there is instruction with paper-and-pencil methods prior to calculator introduction (Hembree & Dessart, 1986, 1992). Other benefits beyond these also are seen. The inclusion of calculators in the classroom and on assessments has been found to positively impact student attitudes. Hopkins (1992) stated, "the presence of the calculator made them [the students] feel more confident, and therefore more positive, about the testing situation" (p. 165). The confidence created by calculator inclusion can be attributed to an accessibility to concepts normally reserved for older students, thereby engendering a feeling of intelligence (Finley, 1992). In addition to better attitudes, provision of a calculator has been found to encourage more careful measurement because the increased precision does not negatively impact the computational process (Reys, 1989).

This body of evidence points to the reason why calculators should be included during instruction and assessment; however, calculator inclusion has some aspects that can be viewed as problematic. One such aspect is the issue of time. Although Hopkins (1992) pointed out that the inclusion of calculators safeguards students from excessive time spent on computation, researchers have documented that an effect of calculator use in standardized testing has been the increased time necessary to complete certain test items, particularly word problems and items that involve computations with fractions (Carpenter, Corbett, Kepner, Lindquist, & Reys, 1981; Lewis & Hoover, 1981). There is no clear evidence as to how much of this extra time is due to the unproductive pushing of buttons or productive explorations into various methods of solving the problem. If students are spending extra time

“playing” or becoming accustomed to the calculator, this could be a major deficiency associated with making calculators available during testing; however, if students are spending time exploring the problem space, the validity of timed tests may need to be evaluated. According to Wilson and Kilpatrick (1989), even if instruction on how to operate a calculator is provided, students unfamiliar with the provided calculator might not have the opportunity to discover how the model aids in the solution of particular problems or how to efficiently chain operations using the calculator’s memory, thereby biasing the results.

Another obstacle associated with calculator inclusion in the classroom and assessment is the amount of time spent on training students to use the calculator correctly. Instead of spending time on learning mathematics, instructional time must be spent familiarizing students with the calculator. The variety of calculators available to students can also pose problems. For example, a student who has a fraction-friendly calculator, that is, a calculator with the capability to present output in fraction form, has an advantage over a student without access to a fraction-friendly calculator on an assignment asking students to record their answers in fraction form. An advanced calculator with programming and graphical capabilities provides a student with opportunities to utilize a wider range of problem-solving techniques than would be available to a student using a four-function or scientific calculator. This technological edge results in an undue advantage for the student with the technological superiority. In addition, the use of programmable calculators raises security issues for standardized testing and tests administered to several classes (Wilson & Kilpatrick, 1989).

COMPONENTS OF CALCULATOR-AVAILABLE ASSESSMENTS

There are three types of calculator-available assessments—calculator based, calculator passive, and calculator neutral—each with its own focus and composition. These tests not only provide information concerning students’ mathematical capabilities, they also can provide information concerning students’ judicious usage of the calculator. A fourth type of assessment, which is discussed in this section, is called *calculator-specific assessment* because it is designed to focus on student mastery of calculator techniques.

Calculator-Based Assessments

Of all the calculator-available assessments, calculator-based tests are designed to gather the greatest amount of information on student understanding of the effective employment of the technology. To achieve this, calculator-based tests are designed

to utilize three types of items—calculator active, calculator inactive, and calculator neutral—and to include a prompt asking students to report calculator usage. The three item types are blended throughout the assessment to examine student performance on items with various calculator requirements and to identify whether students are using the calculator appropriately. In addition, the prompt that is attached to each item permits the assessor to gather information on students’ employment of the calculator. The following discussion further identifies the characteristics of these three item types.

Calculator-active items. The presence of calculator-active items distinguishes the calculator-based assessment from the other calculator-available assessments. Calculator-active items practically require calculator usage because students would find it challenging to solve the questions without the calculator’s aid. Two exemplary items, Figures 1 and 2, were found in the released items from the 12th-grade exam on the 1992 NAEP. The first item, shown in Figure 1, requires students to examine and calculate the price of goods using the two different rate systems. In particular, students must interpret the two graduated rate systems with special attention paid to the phrases “Every 10th tape **free** for one night” and “Every 10 credits = one **free** rental.” For Video Store A, the phrase “Every 10th tape **free** for one night” means that one pays for the first 9 rentals, and the 10th is free. This results in a computation of the price for 20 one-night rentals, 7 two-night rentals, and 3 free 1-night rentals, which is \$82.05. For Video Store B, the phrase “Every 10 credits = one **free** rental” means that the 11th tape is free, and this results in a scenario of 21 one-night rentals, 7 two-night rentals, and 2 free 1-night rentals, a total cost of \$84.00. The computation for this task, although not impossible, may be considered prohibitive in terms of time expenditure. In the item presented in Figure 2, the student is required to estimate the value of N to the nearest tenth. The computational complexity inherent to this item extends beyond simple multiplication and requires the expression of the square root of 8 as a decimal approximation and the use of that approximation to establish a value for N . Thus, calculator-active items practically necessitate the use of a calculator in order to obtain a solution.

Calculator-inactive items. Calculator-inactive items are on the other end of the spectrum from calculator-active items because they are questions for which the calculator is not employable. Questions in this category normally do not require computation but rather require either knowledge retrieval or conceptual answers. The examples from the NAEP (1988) of calculator-inactive items are shown in Figure 3. In the first item, the student must conceptualize which operation would provide the correct solution to the problem situation presented, a process that cannot be aided by the use of a calculator. The second item requires the student to plot the

FIGURE 2 Example of a calculator-active item. Note. From 1992 Released Grade 12 NAEP Calculator Blocks (Accession No. HW000648), by National Assessment of Educational Progress, 1992.

In the equation above, what is the value of N , rounded to the nearest tenth

$$\sqrt[3]{8} N = 3^5$$

Answer: _____

Did you use a calculator on this question

Yes No

FIGURE 1 Example of a calculator-active item. Note. From 1992 Released Grade 12 NAEP Calculator Blocks (Accession No. OM000493), by National Assessment of Educational Progress, 1992.

The Peterson family rents 30 videotapes yearly, of which 23 are rented for one night only and 7 are rented over a period of two nights. Given the rental fee structure shown above, fill in the chart below with the total yearly cost for the Petersons at each store. (Note: The 30 tapes include the free tapes earned.)

Store	A	B
Total Cost		

Did you use a calculator on this question

Yes No

Video Store A
 \$2.65 per tape for one night
 \$1.50 charge for each additional night
 Every 10th tape **free** for one night

Video Store B
 \$3.00 per tape for 2 nights
 1 credit if tape returned after one night
 Every 10 credits = one **free** rental

point (1, 4) on a rectangular grid and assesses whether the student has a specific piece of knowledge and an associated procedure to integrate the understanding into a plotting procedure. Both of these items require students to recall knowledge and conceptualize the provided situation in a context in which calculations are not necessary. Thus, calculator-inactive items tend to be concept oriented with either recall or explanation as the primary focus.

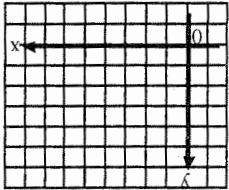
Calculator-neutral items. Calculator-neutral items are those that do not require calculator use to obtain a solution; however, some students might choose to use calculators. The major difference between a calculator-neutral item and a calculator-inactive item is that a calculator-neutral item includes a computation that should be readily obtainable using standard paper-and-pencil techniques, whereas a calculator-inactive item includes no such computation. The examples from the NAEP (1988) of calculator-neutral items are shown in Figure 4. For the first item, the solution process with a calculator would involve multiplying 14 and 28 and then comparing the resultant value to the various foils. Because the item is a two-digit by two-digit multiplication, a student at either the 4th- or 8th-grade level should be able to solve the problem with or without a calculator. The second item requires the student to calculate 3% of 2½ lb in terms of ounces. The student must make a conversion of pounds to ounces using the provided relation. Then, a student not using a calculator possibly would have to convert 3% into 0.03 or 3/100 and then proceed on to the calculation, whereas a student using a calculator could simply calculate 3% × 40 ounces. Either way, the task should be attainable for 8th- and 12th-grade students with or without the use of a calculator.

The blend of calculator-active, calculator-inactive, and calculator-neutral items permits assessors to identify whether students are cognizant of the varying needs for calculator implementation and whether students are proficient and judicious users of the calculator technology. The inclusion of calculator-active items distinguishes calculator-based assessments from either calculator-passive or calculator-neutral assessments. These other two assessments, although permitting calculator use, do not involve complex computations that would be benefited by calculator intervention.

Calculator-Passive Assessments

In contrast to the calculator-based assessments' use of all three item types, calculator-passive tests utilize only calculator-inactive items. As a result, the tests are concept oriented and do not require the use of computational strategies. Such a test is beneficial either when students have not developed proficiency with the calculator and one wishes to avoid student errors associated with improper calculator

FIGURE 3 Example of calculator-inactive items. Note. From *Mathematics Objectives 1990 Assessment* (p. 62), by National Assessment of Educational Progress, 1988, Princeton, NJ: Educational Testing Service.

<p style="text-align: center;">Grades: 8 and 12</p> <p>The number of minutes in 2 days could be determined by performing which of the following operations?</p> <p>(A) Multiplying 60 by 2 (B) Dividing 60 by 24 (C) Multiplying 60 by 24 (D) Dividing 120 by 24 (E) Multiplying 60 by 48</p>	<p style="text-align: center;">Grades: 8 and 12</p> <p>Plot the point (1,4) on the rectangular coordinate grid shown below.</p> 
--	--

use or when the goal of the assessment is focused on the conceptual aspects of the material. Either way, a calculator-passive assessment provides the assessor with no direct information concerning students' calculator capabilities because the only information derivable is whether students realized that the calculator would not be effective on this type of test. Calculator-passive tests are included with this discussion of calculator-available assessments because calculator-passive assessments allow the use of calculators even though the items do not elicit computations directly.

Calculator-Neutral Assessments

Calculator-neutral assessments employ calculator-inactive and calculator-neutral items and do not include any calculator-active items. As a result, calculator-neutral tests do not necessitate the use of the calculator in solving any of its items. This constraint permits an assessor to ease the students into a calculator-available assessment without concern about the calculator negatively interacting with either the difficulty level or the objective of items. At the same time, students are provided with an opportunity to use both calculator and by-hand strategies. One should be aware that calculator-neutral assessments are not useful measures of students' effective employment of the technology because the items being used do not require calculator use to be successful. Thus, the only information obtained regarding students' calculator capabilities is whether students knew to avoid using the calculator on the calculator-inactive items and whether they were able to choose a reasonable strategy, either using a calculator or by hand, for answering the calculator-neutral items.

Calculator-Specific Assessments

Calculator-specific tests differ from the three calculator-available assessments previously described because they focus on ascertaining student proficiency with the calculator. The items incorporated into this type of test focus on measuring students' recognition of appropriate key sequences, interpretation of calculator display information, and placement and use of specific calculator keys. Examples of items that focus on these various skills are shown in Figures 5 through 7. In the item presented in Figure 5, students are asked to translate a nested computation into the proper key sequence to determine the answer. Here the focus of the item is to see whether students recognize, from the given set of choices, the sequence of key strokes that will attain the answer to the complex calculation. In particular, students must recognize how to use parentheses and to employ the exponential key in solving embedded computations. For the item presented in Figure 6, students must make

FIGURE 4 Examples of calculator-neutral items. Note. From *Mathematics Objectives 1990 Assessment* (p. 63), by National Assessment of Educational Progress, 1988, Princeton, NJ: Educational Testing Service.

<p style="text-align: center;">Grade: 4 and 8</p> <p>Multiply</p> $\begin{array}{r} 14 \\ \times 28 \\ \hline \end{array}$ <p>(A) 42 (B) 362 (C) 372 (D) 392 (E) 522</p>	<p style="text-align: center;">Grades: 8 and 12</p> <p>The fat content of a certain food product is known to be approximately 3 percent by weight. Approximately how many ounces of fat are in 2 1/2 pounds of this product? (16 ounces = 1 pound)</p> <p>(A) 0.075 (B) 0.12 (C) 1.2 (D) 7.5 (E) 12.0</p>
---	--

FIGURE 6 Calculator-specific item (2). Note. From *Mathematics '90: A Status Report on School Mathematics in British Columbia—Provincial Report* (p. 32), by D. F. Robitaille, T. L. Schroeder, and C. C. Nicol. Copyright 1991 by the Province of British Columbia. Reprinted with permission.

Someone has just multiplied 6540 by 987 600 on the calculator. The calculator with the answer displayed is shown below. The product of the two numbers is

6.458904 09											
C	CE	ON/OFF	M+	M-	M \times	M \div	%	$\frac{1}{x}$	\sqrt{x}	7	8
4	5	6	\times	7	8	9	-	0	.	$\frac{1}{x}$	=
1	2	3	\div	4	5	6	\times	7	8	9	-
0	.	$\frac{1}{x}$	=	1	2	3	\div	4	5	6	\times

A. 6 458 904 000
 B. 6 458 904 009
 C. 6 458 9049
 D. 645 890 400
 E. I don't know.

Which calculator key sequence would you use to obtain the answer to

$$\sqrt[3]{\frac{(2+6^3)(5-7)^2}{5}}=?$$

A. 2 + 6 \sqrt{x} 3 \times 5 - 7 \sqrt{x} 2 \div 5 \sqrt{x} 3
 B. 6 \sqrt{x} 3 + 2 \times 5 - 7 \sqrt{x} 2 \sqrt{x} 3 \div 5
 C. (2 + 6 \sqrt{x} 3) \times (5 - 7 \sqrt{x} 2) \div 5 = \sqrt{x} 3
 D. 6 \sqrt{x} 3 + 2 = \times (5 - 7) \sqrt{x} 2 \div 5 = \sqrt{x} 3

FIGURE 5 Calculator-specific item (1).

sense of the calculator display. The focus of this item is the interpretation of the cryptic scientific notation presented. Students need to know how to interpret and translate the scientific form of the presented number. The item shown in Figure 7 could be used to determine whether students know where particular keys can be found on the calculator. In general, finding the square root of 29 requires the use of a calculator, especially when the request is for seven-place accuracy. If students are unable to locate the square root button on their calculators, obtaining a reasonable approximation is unlikely.

The Mathematical Association of America and the College Board oppose placing any question focused on calculator skills on any standardized test (Kenelly, 1989); however, calculator-specific tests have been found to be useful in determining a student's proficiency in using and interpreting displayed results. In particular, calculator-specific tests can be utilized to ascertain whether students have developed mastery of the technological tool prior to engaging in assessments that might require the use of the calculator.

ADDITIONAL TEST DEVELOPMENT CONSIDERATIONS

In this section, the focus turns to some additional considerations that must be addressed when designing any calculator-available assessment. The first part focuses on test construction, and the second part discusses several analyses, at the item level, that must be performed prior to the administration of any calculator-available test.

Test Construction Considerations

Developing calculator-available assessments can be accomplished in two ways. The first assumes the development of an entire assessment from the selection of items pooled into the categories calculator active, calculator inactive, or calculator neutral. Depending on the types of items selected, a calculator-based, calculator-passive, or calculator-neutral assessment can be constructed. The other method of fabricating a calculator-available assessment is called retrofitting. These two methods of test construction are now discussed.

Item pools. Calculator-based, calculator-passive, and calculator-neutral assessments can be built from pools of items in which the items have been designed to be of the three primary categories: calculator active, calculator inactive, or calculator neutral. Once items have been designed and categorized, the items are selected to be included in the calculator-available assessment. Depending on the

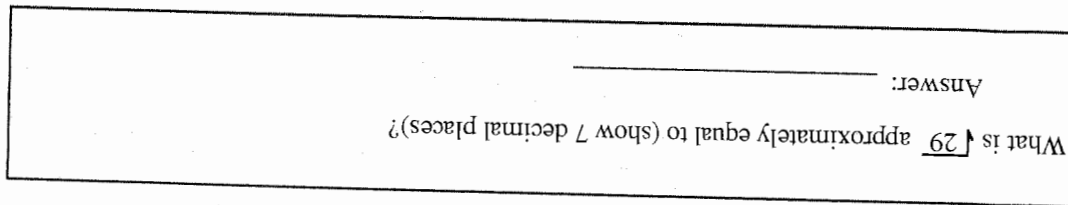


FIGURE 7 Calculator-specific item (3).

type of calculator-available assessment desired, the composition of item types in the assessment must be adjusted. In particular, any item categorized as calculator active, calculator inactive, or calculator neutral can be used on a calculator-based test with the stipulation that at least some calculator-active items be included. Calculator-neutral assessments permit the combination of calculator-inactive and calculator-neutral items with the stipulation that at least some calculator-neutral items be included. The least diverse assessments with respect to the item composition are the calculator-passive tests because they are composed of only calculator-inactive items.

Retrofitting. The retrofitting method takes an assessment not designed for the allowance of calculator technology, examines the various items, removes items inconsistent with the type of calculator-available assessment desired, and replaces them with items that are consistent. Retrofitting an assessment to obtain a calculator-based test requires the addition of calculator-active items, which necessitate the use of the calculator. On the other hand, a calculator-passive assessment can easily be developed by removing any items that require computation, that is, calculator-active or calculator-neutral items, and replacing them with calculator-inactive items focused on knowledge retrieval or conceptual answers. The design of calculator-neutral assessments using retrofitting requires more analysis. To obtain a calculator-neutral test, it is necessary to remove any items that would benefit calculator users, that is, calculator-active items. To aid in distinguishing which items have to be removed from a potential calculator-neutral exam, Lietzel and Waits (1989) developed a calculator-sensitivity (CS) index: "CS = percentage correct by calculator-using group/percentage correct by non-calculator-using group" (p. 22). The higher the CS value for a particular item, the greater the disparity between students using calculators and students not using calculators on the item. This implies that the item favors the use of the calculator and should be removed. Ideally, a calculator-neutral test should not advantage either group of students. Thus, if an item on a calculator-neutral test has a CS value significantly above or below 1.00, it should be removed from the test. Once all items of this type have been removed, they can be replaced with either calculator-inactive items, for which the calculator would not be beneficial, or calculator-neutral items, which are computational but do not mandate calculator use. The primary problem associated with establishing the CS index for any item or set of items to be used on a test is that the items must be pilot tested with two comparable groups of students: one group permitted to use calculators and the other group not permitted to use calculators.

Although these two methods of test construction produce calculator-available assessments that appear to be coherent at the global level, item-level problems can exist, especially when items have not been pilot tested. The next section outlines additional points that must be considered in developing calculator-available tests.

Item-Level Considerations

Making the calculator an available option can create difficulties at the item level. For instance, inappropriate questions may exist on a test, resulting in students' scores that are not reflective of their understanding of the material but rather of the tool. In addition, on a multiple-choice assessment, the nature of the foils can potentially pose problems. If the foils do not correlate well with calculator use, the results of the assessment may not be indicative of student knowledge because of errors in translation. A movement to the use of open-ended assessments will not alleviate all the problems. Because the goal of an open-ended assessment is to have students provide both product and process, and because much, if not all, of the computational work will reside in the calculator, there are difficulties in having students provide sufficient material to reveal the process employed. Each of these problems is discussed in the ensuing paragraphs.

Looking for inappropriate questions. The use of a test without inspection of the items can cause difficulties. It might be hazardous to permit students to use a calculator on just any mathematics assessment because the test might include inappropriate, or "calculator-unacceptable," items (Harvey, 1992). An item can be considered unacceptable for inclusion on a calculator-available assessment if the item's objective or difficulty level changes according to calculator use. Some examples of unacceptable questions are shown in Figure 8. The first item's difficulty level changes when a calculator is allowed. In particular, if a student has a fraction-friendly calculator such as the TI-Math Explorer available, the task simply becomes one of correctly translating the stem into key punches because the calculator's output would be in the form of a reduced fraction. Even if the student has only a four- or five-function calculator available, the difficulty of computing with a fraction in the denominator is easily overcome by the calculator's conversion into decimals. With respect to the second item, the objective changes from testing understanding of fractional exponents to key-sequencing the stem. If a scientific calculator were used by a student solving the third item, both the difficulty level and the objective would change. The task would simply become one of testing each of the presented foils.

Another type of inappropriate item, which can appear in open-ended or multiple-choice exams, is one that provides students the key sequence used to solve the problem. For example, verbal clues such as "divide a by b " suggest an input ordering, thus increasing student performance (Kouba & Swafford, 1989). This type of task becomes extremely trivial to a student using a calculator, therefore giving that student an undue advantage over a student who is not using a calculator.

A final type of inappropriate question is one that appears to ask for one thing but requires another. The example provided in Figure 9 illustrates how placing an

FIGURE 9 Inappropriate National Assessment of Educational Progress fourth-grade calculator item. Note. From *Can Students Do Mathematical Problem-Solving: Results From Constructed-Response Questions in NAEP's 1992 Mathematical Assessment*, by J. A. Dossey, I. V. S. Mullis, and C. O. Jones, 1994, Washington, DC: National Center for Educational Statistics.

Think carefully about the following question. Write a complete answer. You may use drawings, words and numbers to explain your answer. Be sure to show all of your work.

Laura wanted to enter the number 8375 into her calculator. By mistake, she entered the number 8275. Without clearing the calculator, how could she correct her mistake?

Without clearing the calculator, how could she correct her mistake another way?

FIGURE 8 Calculator-unacceptable tasks. Note. Released items on the MAA Committee's Placement Examinations Mathematics Tests CR/1B and A/4A. Mathematical Association of America © 1992. Reprinted with permission.

1. $\frac{2}{2+\frac{1}{3}} =$
 - (a) $\frac{4}{3}$
 - (b) $\frac{7}{6}$
 - (c) $\frac{3}{4}$
 - (d) 2
 - (e) 3
2. $(8^{-1/3})(9^{1/2}) =$
 - (a) 6
 - (b) -6
 - (c) $(72)^{-1/2}$
 - (d) $2/3$
 - (e) $3/2$
3. For which values of x is $\tan x$ not defined?
 - (a) $-\pi$
 - (b) $-\pi/2$
 - (c) 0
 - (d) $\pi/4$
 - (e) $\pi/3$

item in a context in which calculators can be employed may mislead students. In the item presented in Figure 9, it is difficult to ascertain if place value, addition and subtraction of 100, or calculator use is being assessed. Examination of the item's associated scoring guide reveals that the designers sought to assess knowledge of place value (i.e., recognition that the difference between 8,375 and 8,275 occurs in the hundreds place) rather than calculator proficiency or ingenuity (i.e., correction of an entry error without employing the On/C key; Silver & Kenney, 1993). The scoring guide appeared to discriminate between those who applied a mathematical method to make the proper adjustment and those who applied a "clever calculator trick" (Silver & Kenney, 1993, p. 23). In particular, Silver and Kenney indicated that the availability of four-function calculators was misleading on two accounts because

the scoring appeared to penalize unfairly those students who focused on calculator use instead of on place value but also ... the presence of the calculator may actually have been somewhat detrimental rather than helpful because it may have distracted students' attention from the mathematical concepts at the heart of the item. (p. 23)

As a result of such discrimination, those who employed a mathematically based strategy would score better than would students who focused on calculator use. Thus, if the focus of the item is only on mathematically based strategies, it would seem imprudent to situate the context with respect to calculator use because students may be misled into focusing on calculator-based strategies. As a result, the items of a calculator-available assessment must be examined to ensure that calculator usage does not affect an item's objective and that the scoring guide is consistent with calculator strategies when calculators are permitted.

Foils of multiple-choice questions. Another consideration that must be addressed when employing a multiple-choice assessment instrument with the inclusion of the calculator is the nature of the foils. For example, "a multiple-choice algebra test that gives all answers in radical form, for example, $(2 + \sqrt{3})/4$, might not be as appropriate for students using calculators as a test that includes answers that are decimal approximations" (NCTM, 1989, p. 195). If a student were using a calculator to solve the problem, it would be necessary for the student to obtain decimal approximations of the foils, determine the answer for the stem, and make a comparison between the answer obtained and the values calculated for the foils, thereby increasing the amount of time necessary to solve the item. One might ask the question, if students are presented with foils in radical form, why would they use calculators? Any response to this question must first address the purpose of the item. If students' algebraic knowledge is being assessed, it would not matter if the foils were in radical form or decimal expressions. If the purpose of the assessment

is to determine whether students know how to obtain exact solutions to the algebraic problem, it would be inappropriate for students to employ the technology because it is less efficient. Finally, if the goal of the assessment is to determine whether students know when it is practical to employ technology, it is quite artificial to use the foils to mandate calculator use. In any case, the matter of the foils must be addressed in light of the goals of the item and the assessment and how the item reflects those goals.

Prompts of open-ended questions. The problems associated with foils are overcome by the openness of open-ended assessments. However, when students use calculators on open-ended assessments, another difficulty is encountered. For a multiple-choice test, the information the student places on the paper is a chosen foil, whereas open-ended assessments require students to provide both the process and the product. If an open-ended task requests a student to "show your work," the student may believe there is nothing to write down because all the work was done in the calculator. As a result, the prompts that accompany open-ended assessments in a calculator-available atmosphere should be framed more carefully. Moskal and Meel (1993, p. 11) identified alternative prompts ("Explain your answer"; "Show how you found your answer. You should include your calculations and give the results of your calculations"; "Show how you found your answer. You should include all the steps you completed to find your answer"; and "Show your work. Your work should be clear enough so that another person could read it and understand your thinking") that elicited responses that were more robust than were those from prompts such as "Show your work" or "Show how you found your answer" and provided the assessor with information concerning a student's process and product. Two examples are included in Figure 10.

In Figure 10, one can see that the student responding to the more directive prompt provided more information than did the student responding to the less directive prompt. The more extensive response permits the assessor to see the student's work rather than have it reside in the calculator. The NAEP provides students with the following prompt on their open-ended questions when calculators are available:

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show *all* your work. (NAEP, 1992)

This level of specificity implores students to provide a written account of computations even if the calculator was used. Thus, the use of calculator-available assessments necessitates the use of more directive prompts calling students to provide a record of the computations done with the calculator.

CONCLUSION AND RECOMMENDATIONS

The inclusion of calculators in the classroom and in the domain of assessment offers a considerable number of benefits. Students have a better attitude about the testing situation and feel empowered (Bitter & Hatfield, 1992; Finley, 1992; Hembree & Dessart, 1986, 1992; Hopkins, 1992; Mungler & Loyd, 1989; G. Wheatley & Shumway, 1992). They are freed by the de-emphasis of computation to attend to more important aspects of the mathematics curriculum, such as problem solving, reasoning, and conceptual development, without diminishing computational abilities when accompanied with previous work on paper-and-pencil methods (Hopkins, 1992; Vanatta & Hutton, 1980; C. L. Wheatley, 1980; Wilson & Kilpatrick, 1989). They are able to engage in problem-solving activities with realistic tasks relevant to students rather than with contrived problems (Hopkins, 1992; Wilson & Kilpatrick, 1989). With the aid of calculators, students can encounter complex mathematical topics at younger ages than previously thought possible. For example, the exploration of maxima and minima, which previously had been reserved for students studying algebra, can now be introduced to middle school students with graphing calculators during a study of functions. When students are allowed to use calculators, they become skillful in determining when technology would be beneficial for solving problems.

However, complications may confront teachers and test developers in implementing calculators in both classrooms and assessments. If calculators are used in a timed assessment, students might spend more time on particular items and be unable to complete the assessment. Another difficulty is that different calculators can give some students undue advantages. The nature of particular items can cause difficulties. Inappropriate questions may exist on a test, resulting in students' scores that are reflective of their understanding of the tool rather than of the material to be assessed. Multiple-choice foils present problems in calculator-available assessments. If foils do not correlate well with calculator use, the results of the assessment may not be indicative of student knowledge because of errors in translation. Open-ended questions in a calculator-available test are not without drawbacks. Much, if not all, of the computational work will reside in the calculator, reducing the amount of information provided by the student on the paper. As a result, explicit prompts must be developed to ensure valid scoring of responses.

Several recommendations can be made as part of this conclusion. In particular, the following set of guiding principles should be considered when calculator-available assessments are being implemented:

1. The constraints on time should be taken into account in implementing calculator-available assessments. Students may need additional time depending on the real-world complexity of the tasks comprising the assessment.

Randall likes to go skating. The signs in the front of the rink say:

<p>Members \$7.00 at the beginning of the season Plus \$0.50 to rent skates each visit</p>	<p>Non-Members \$1.50 for each visit This includes skate rental.</p>
--	---

Randall does not own skates so he must rent them.

How many times would Randall have to go skating, so that being a member would cost the same amount as being a non-member?

General Prompt:

Show your work.

Information Received:

Answer: 7

Did you use the calculator on this question? yes no

Directive Prompt:

Show your work. Your work should be clear enough so that another person could read it and understand your thinking.

Information received:

MEMBERS - EACH VISIT HOW MANY TIMES
 $\$1.50 \times 7 = \10.50

NON-MEMBERS - HOW MANY TIMES SKATE RENTAL
 $7 \times .50 = \$3.50$
 $7.00 + 3.50 = 10.50$

Answer: 7

Did you use the calculator on this question? yes no

FIGURE 10 Student responses to the skating task with two alternative prompts.

2. The set of available calculators should not unduly advantage one group of students over others.
3. The items of a calculator-passive test should be inspected to ensure no computational items are present.
4. The items of a calculator-neutral test should focus on concepts, ideas, and calculations that can easily be attained by hand or by calculator. Inspection of the items prior to administration should identify and remove any items that necessitate calculator use.
5. The items of a calculator-based test must include calculator-active items practically necessitating calculator use. A calculator-based test should not be used unless students are calculator proficient; a calculator-specific test is helpful in determining this. In addition, the inclusion of a prompt asking students to report calculator usage at the item level is beneficial in gathering information concerning students' appropriate employment of the calculator.
6. The test items should be examined to determine whether the difficulty level or objective changes when a calculator is used. If either the difficulty level or objective changes, the assessor should decide whether the change is appropriate for the goals of the assessment.
7. The test items should be examined to determine if the stem provides the key sequence to solve the problem. If an item's stem does provide the key sequence, then what is being assessed from a calculator-using student is only the ability to key-in the stem and read the answer.
8. The test items should be examined to determine whether the problem situation coupled with the calculator availability inadvertently misleads students into focusing on calculator processes rather than mathematically based strategies. If mathematically based responses are considered the only valid responses, either the scoring guide should be adapted to allow for calculator-based responses or the task should be modified to eliminate the misleading aspects.
9. The multiple-choice items of an assessment should be examined to ensure that the foils are consistent with calculator use, the goals of the assessment, or both.
10. The open-ended items of an assessment should be examined to ensure that the prompts induce students to provide sufficient information on the test paper rather than allow it to reside in the calculator's memory.

These recommendations should aid in overcoming some of the problematic areas associated with designing and implementing calculator-available assessments. However, one should remember that the inclusion of the calculator in the testing situation is not a panacea. The calculator is merely a tool to be embedded within the goals of the classroom and assessment, providing students with an opportunity in the classroom to develop some of the skills needed later in life.

Correspondence concerning this article should be sent to my current affiliation, the Department of Mathematics and Statistics, Bowling Green State University.

An earlier version of this article was presented at the April 1995 meeting of the NCTM in Boston, Massachusetts.

REFERENCES

- ACT allows calculators. (1996, September). *NCTM News Bulletin*, p. 12.
- Bitter, G. G., & Hatfield, M. M. (1992). Implementing calculator in middle school mathematics: Impact on teaching and learning. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 200–207). Reston, VA: National Council of Teachers of Mathematics.
- Braswell, J. S. (1992). Changes in the SAT in 1994. *The Mathematics Teacher*, 85, 16–21.
- Brüggeman, B., Harvey, A., & Braswell, J. (1993, April). *Effects of calculator use on SAT-M performance*. Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.
- Carpenter, T. P., Corbett, M. K., Kepner, H. S., Lindquist, M. M., & Reys, R. E. (1981). Calculators in testing situations: Results and implications from national assessment. *Arithmetic Teacher*, 28(5), 34–37.
- Dossey, J. A., Mullis, I. V. S., & Jones, C. O. (1994). *Can students do mathematical problem solving: Results from constructed-response questions in NAEP's 1992 Mathematical Assessment*. Washington, DC: National Center for Education Statistics.
- Finley, K. W. (1992). Calculators add up to math magic in the classroom. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 195–199). Reston, VA: National Council of Teachers of Mathematics.
- Greenes, C. E., & Rigol, G. W. (1992). The use of calculators on college board standardized tests. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 187–194). Reston, VA: National Council of Teachers of Mathematics.
- Harvey, J. G. (1992). Mathematics testing with calculators: Ransoming the hostages. In T. Romberg (Ed.), *Mathematical assessment and evaluation: Imperatives for mathematics educators* (pp. 139–168). Albany: State University of New York Press.
- Hembree, R., & Dessart, D. J. (1986). Effects of hand-held calculators in precollege mathematics education: A meta-analysis. *Journal for Research in Mathematics Education*, 17, 83–99.
- Hembree, R., & Dessart, D. (1992). Research on calculators in mathematics education. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 23–32). Reston, VA: National Council of Teachers of Mathematics.
- Hopkins, M. H. (1992). The use of calculators in the assessment of mathematics achievement. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 158–166). Reston, VA: National Council of Teachers of Mathematics.
- Kenelly, J. W. (1989). A historical perspective on calculator usage on standardized examinations in mathematics. In J. Kenelly (Ed.), *The use of calculators in the standardized testing of mathematics* (pp. 7–15). New York: College Entrance Examination Board.
- Kouba, V., & Swafford, J. (1989). Calculators. In M. Lindquist (Ed.), *Results from the Fourth Mathematics Assessment of the National Assessment of Educational Progress* (pp. 94–105). Reston, VA: National Council of Teachers of Mathematics.

ACKNOWLEDGMENTS

Dynamic Assessment of Composing Abilities in Children With Learning Disabilities

Grant Cioffi and John J. Carney
University of New Hampshire

Dynamic assessment shifts the perspective in educational diagnosis from measuring products to understanding processes. From this perspective, dynamic assessment offers a framework for the educational diagnosis of writing disabilities that is consistent with process models of writing instruction. The procedures employ interview and response-to-instruction paradigms to study both children's writing abilities and their knowledge of procedures for writing. The diagnostic procedures sample children's proficiency with, and knowledge of, planning, drafting, and revising. Diagnostic protocols are outlined, and guidelines for the interpretation and formulation of instructional recommendations are offered.

In schools, clinicians and teachers routinely must engage in diagnostic assessment. In an ideal world, teaching and development would be inextricably linked. As teachers observed children's intellectual, social, and academic growth, they also would seek to foster such progress. Routine observations with respect to children's abilities would be complemented with consideration of how instruction or the classroom environment supports learning. This activity, the close observation of learning in response to instruction, is an exercise in problem solving and constitutes an essential element of diagnostic assessment. Whether in the classroom or the clinic, the interest is not simply to categorize a child under a particular label or diagnostic scheme but also to uncover the kind of instruction and educational setting that will lead to continued development. More specifically, the teacher must construct (a) an understanding of the child's current abilities, (b) a description of the kind of instruction the child should receive to make progress, and (c) a recommendation regarding the extent and intensity of intervention that is necessary.

Requests for reprints should be sent to Grant Cioffi, Department of Education, Morrill Hall, University of New Hampshire, 62 College Road, Durham, NH 03824-3595.

- Lewis, J. C., & Hoover, H. D. (1981, April). *The effect on pupil performance from using hand-held calculators during standardized mathematics achievement tests*. Paper presented at the meeting of the American Educational Research Association, Los Angeles.
- Lienwand, S. J. (1992). Calculator in state testing: A case study. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 167-176). Reston, VA: National Council of Teachers of Mathematics.
- Lietzel, J. R., & Waits, B. K. (1989). The effects of calculator use on course tests and on statewide mathematics placement tests. In J. Kenelly (Ed.), *The use of calculators in the standardized testing of mathematics* (pp. 7-24). New York: College Entrance Examination Board.
- Loyd, B. H. (1991). Mathematics test performance: The effects of item type and calculator use. *Applied Measurement in Education*, 4, 11-22.
- Moskal, B., & Meel, D. E. (1993). *Calculators and performance assessments: Issues related to the development of tasks*. Unpublished manuscript, University of Pittsburgh, Pittsburgh, PA.
- Mullis, I., Dossey, J., Owen, E., & Phillips, G. (1991). *The state of mathematics achievement: NAEP's 1990 Assessment of the Nation and the Trial Assessment of the States*. Washington, DC: U.S. Government Printing Office.
- Munger, G. F., & Loyd, B. H. (1989). Gender and attitudes toward computers and calculators: Their relationship to math performance. *Journal of Educational Computing Research*, 5, 167-177.
- National Assessment of Educational Progress. (1988). *Mathematics objectives 1990 assessment*. Princeton, NJ: Educational Testing Service.
- National Assessment of Educational Progress. (1992). *1992 released Grade 12 NAEP calculator blocks*. (Available from the Educational Testing Service, Rosedale Road, Princeton, NJ 08541-0001)
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Pandey, T. N. (1986, March). *The influence of testing on mathematics education*. Paper presented at the national conference on the Influence of Testing on Mathematics Education, Los Angeles.
- Payne, J. N. (1992). The effects of calculator on state objectives and test. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 177-185). Reston, VA: National Council of Teachers of Mathematics.
- Reys, B. J. (1989). The calculator as a tool for instruction and learning. In J. Fey (Ed.), *New directions for elementary school mathematics—Yearbook* (pp. 168-173). Reston, VA: National Council of Teachers of Mathematics.
- Robitaille, D. F., Schroeder, T. L., & Nicol, C. C. (1991). *Mathematics '90: A status report on school mathematics in British Columbia—Provincial report*. Victoria, British Columbia, Canada: Ministry of Education and Ministry Responsible for Multiculturalism and Human Rights.
- Silver, E. A., & Kenney, P. (1993). *The content and curricular validity of the 1992 NAEP TSA in mathematics*. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.
- Silverman, F. L., & Muehl, S. A. (1989). Calculators: On balance, they will enrich school mathematics. *Illinois Schools Journal*, 69(1), 29-35.
- Vanatta, G. D., & Hutton, L. A. (1980). A case for the calculator. *Arithmetic Teacher*, 27(9), 30-31.
- Wheatley, C. L. (1980). Calculator use and problem-solving performance. *Journal for Research in Mathematics Education*, 11, 323-334.
- Wheatley, G., & Shumway, R. (1992). The potential for calculators to transform elementary school mathematics. In J. Fey (Ed.), *Calculators in mathematics education—Yearbook* (pp. 1-8). Reston, VA: National Council of Teachers of Mathematics.
- Wilson, J. W., & Kilpatrick, J. (1989). Theoretical issues in the development of calculator-based mathematics tests. In J. Kenelly (Ed.), *The use of calculators in the standardized testing of mathematics* (pp. 7-15). New York: College Entrance Examination Board.