Assessment Issues Involving Graphics Calculators

Barry Kissane

Australian Institute of Education, Murdoch University, Western Australia

Why the concern for assessment?

Of initial interest is the observation that a substantial portion of the meeting time of this conference has been devoted to issues related to assessment. At first sight, this ought to be surprising. Indeed, my own view is that a graphics calculator is mainly a device to support the learning of mathematics. It has the potential to be important for learning because it can provide access to particular kinds of educational and mathematical experiences (... some of which involve ‘graphing’ and many of which do not, suggesting that the term ‘graphing calculator’ is unfortunate, because of its likelihood of reinforcing the misconception that graphics calculators are mainly useful for graphing). Unlike most other technologies used in schools, graphics calculators were designed for educational purposes. A graphics calculator is a powerful learning device, less powerful as a teaching device and not obviously particularly useful as an assessment device. From this perspective, the degree of emphasis on assessment would seem at first to be unwarranted.

However, the emphasis on assessment within this conference is a clear reflection of the importance of assessment to thinking about almost any issues in mathematics education. Indeed, the links between teaching, learning and assessment are always recognised officially to be important, reflected in the design of official curriculum documentation at various levels. More than a decade ago, the AAMT’s Discussion Paper on Assessment and Reporting in School Mathematics (1988) observed the increasing significance of technology for mathematics education, suggesting that there were significant consequences of these changes for assessment:

Our conception of what constitutes an adequate mathematical education has undergone considerable recent change, resulting in modifications both to the content and to the teaching methodologies of school mathematics. One such change is an increased emphasis on mathematical processes, as distinct from mathematical content. ... The appearance of calculators and microcomputers as tools for both learning and for doing mathematics have changed rapidly and permanently our vision of an adequate mathematical education (p. 2).

Not surprisingly, the AAMT’s Statement on the Use of Calculators and Computers for Mathematics in Australian Schools (1996), a revision of the earlier statement of a decade before, reflected the main reason for attending to the links between technology and assessment:
Assessment practice should reflect good teaching practice. The use of technological resources as integral aids to learning assumes their inclusion in the assessment process. New approaches to assessment will be required at all levels to better reflect the realities of learning within a technological society (p. 5).

The significance of assessment in the minds of many students and teachers can hardly be overstated. It has long been clear that what is assessed in mathematics has a good deal to do with how people interpret what is important and valued. The student question, 'Will this be on the test?', or variations of it, is commonly heard at all levels of school mathematics education — and no less common at the early undergraduate levels as well.

The nature of and conditions for assessment are widely seen to have significant flow-on effects to classrooms. For example, in those states of Australia in which high stakes external examinations are used, assessment policies regarding the use of graphics calculators have enormous influence over the daily realities of classrooms. Thus, in Western Australia, where senior secondary school examinations are conducted on the assumption that students will have access to graphics calculators, most students have a graphics calculator with them for most of the time in mathematics classes. The calculators that are permitted for use in external examinations are also permitted in classroom tests, school examinations and everyday schoolwork. Of course, the acceptance of graphics calculators for assessment is not the only factor here: there has been some re-design of previous courses to accommodate this sort of technology.

In contrast, within states with high stakes external examinations that prohibit the use of graphics calculators (such as New South Wales), graphics calculators are much less evident. It is rather unusual to find students with their own graphics calculator, or even teachers with much active professional interest in this sort of technology. This may be expected to start to change, of course, with the new General Mathematics courses in Years 11 and 12, with talk of graphics calculator use in the relevant HSC examinations at the end of 2001.

As a third example, the state of Queensland is interesting. Without a high stakes external examination, mathematics teachers have more local control over their curriculum. Consequently, there is more diversity of practice evident. In some schools, students have considerable personal access to (and high levels of ownership of) graphics calculators, which have influenced a significant part if the curriculum. In other schools, graphics calculators are not an important part of the curriculum, presumably because of the interests and expertise of the teaching staff at those schools. Thus, while Queensland enjoys some of the advantages of local curriculum ownership, it does not enjoy the advantage of central control.

In each of these three states (and, indeed, in each of the other states and territories in Australia), anyone interested in the role and significance of graphics calculators for mathematics education is necessarily also obliged to consider the matter of assessment in order to understand present practice and predict likely future directions with any measure of confidence.
What kinds of assessment are relevant?

In many casual conversations about assessment, the relationships between assessment and grading are over-emphasised and the importance of assessment as a means of obtaining and providing feedback is under-emphasised. Indeed, many students (and even teachers and parents) seem to confuse assessment with grading and credentialing issues. Frequently, too, assessment in mathematics focuses on formal assessment, and on testing in particular.

Assessment can and should comprise more than formal assessment, involving testing, for purposes of grading. This is no more or less true in the context of graphics calculators than it is more generally. The AAMT Discussion Paper (1988) presented a continuum for assessment, an extract of which is shown below.

<table>
<thead>
<tr>
<th>Less structured</th>
<th>More structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation</td>
<td>Examinations</td>
</tr>
<tr>
<td>Observation</td>
<td>Classroom</td>
</tr>
<tr>
<td>schedule</td>
<td>checklist</td>
</tr>
<tr>
<td>Student</td>
<td>Projects</td>
</tr>
<tr>
<td>self-assessment</td>
<td>Assignments</td>
</tr>
</tbody>
</table>

Such a continuum suggests some of the range of assessment practices that need to be considered with graphics calculators in mind. It is salutary to be reminded that, while more formal means of assessment dominate the discussion and thinking of many students and their teachers, less formal methods, such as various observational methods, may in practice provide more useful feedback upon which suitable educational plans can be built. That is, what is most publicly valued is not necessarily what teachers and students will find most useful.

David Clarke (1997) observed that 'teachers are assessing all the time':

> Historically, assessment has been seen as a separate activity from instruction, and only forms of activity that involved written student product under testlike conditions have been sanctified with the label 'Assessment'. An even more restrictive view of assessment tends to associate assessment only with activities that lead to the grading of student performance. In short, we have tended to label as 'assessment' only a certain subset of our information gathering and exchanging. Is it any wonder that so many students have come to prize earning a grade over genuine learning — or even to confuse the two?

> A teacher observing a student at work is assessing.

> A teacher engaging in a class discussion is assessing.

> A teacher talking to a student about his or her performance is assessing (p. 20).

In similar vein, the excellent MCIP Assessment Alternatives in Mathematics package (Clarke, 1989) highlighted the many forms of assessment tasks: short answer, open-ended, extended-answers, challenging problems, investigations and investigative projects.
In principle, then, assessment associated with graphics calculators can be oral or written, external or internal to the school and more or less formal. However, formal assessment often carries more weight for grading, is often regarded as more credible, is usually undertaken under controlled conditions such as timing, facilities and communication with other people. Although such assessment is possibly more reliable, since the conditions can be readily replicated, it is probably a less valid way of obtaining important information about student learning than is less formal assessment. Given its public prominence, however, it is necessary to consider formal assessment in detail, with the relevance of graphics calculators in mind.

Issues in formal assessment

Despite much good professional advice, tests and examinations continue to be very important in the Australian educational culture, both within classrooms and external to classrooms. As noted above, the flow-on effects of these are most obvious, since school assessment is usually strongly influenced by external assessment, especially when issues of university entrance are involved (a high stakes situation). However, the flow-on effects of the ways in which graphics calculators are used in formal assessment with a more low stakes character (such as classroom tests) are also likely to be considerable. Hence, this section of the paper focuses on issues for formal assessment.

Many of these issues have been previously described and elaborated in some key papers written by Jen Bradley, Marian Kemp and myself, reflecting philosophical, practical and analytic perspectives. Interested readers are referred to the details of these. The papers were written in the context of the integrated use of graphics calculators in mathematics education, and derived from our experiences at Murdoch University and elsewhere.

Levels of calculator use

A key issue concerns whether or not graphics calculators are used in formal assessment at all, discussed at length by Kissane, Bradley & Kemp (1994). There is a continuum of possibilities, all of which are used at present in parts of Australia:

- Unrestricted use of any calculators
- Unrestricted use of a restricted set of calculators
- Calculator-neutral assessment
- Calculator-free assessment (for some assessment components)
- No calculator use permitted

Another way of thinking about this is to identify three possible choices to be made for calculator use in examinations:

- **Required**
  - Access and instruction are assumed.
  - Students are expected to use where suitable.
  - Student choice is to be exercised — and assessed.
Allowed

(Perhaps begrudgingly?) Students can use calculators if they want to do so.

It is assumed that some students will not have access.

Disallowed

Graphics calculator use is forbidden.

This category has a strong suggestion of control (and gate-keeping?) associated with it.

These issues are discussed in some detail in Kissane, Bradley & Kemp (1994), so that it is not appropriate to repeat the arguments here. Rather, it ought be noted that the arguments for calculator-neutral testing are difficult to make convincingly, in part because it is hard to comprehend that a calculator-neutral test is actually possible without significant sacrifice of content validity, especially given the nature of most senior secondary mathematics courses in Australia today. Both calculator-neutral and calculator-free testing come perilously close to sacrificing the earlier principles suggested of ensuring a measure of coherence between teaching, learning and assessment.

A typology

Based on an analysis of practice, Kemp, Kissane & Bradley (1996) suggested a typology of student uses of graphics calculators in examinations. Three possibilities are identified:

• calculator use is expected

• calculator is used by some students, but not by others

• calculator use is not expected.

Although first constructed some years ago, this typology still seems to be of value today, both at external examination level and within courses; our experience extends to both of these settings.

There are three possible situations in which graphics calculators are expected to be used in examinations:

• students are explicitly advised or even told to use graphics calculators;

• alternatives to graphics calculator use are very inefficient;

• graphics calculators are used as scientific calculators only.

To illustrate but one of these cases, consider an examination task requiring students to solve the equation, \( x^3 = x + 4 \). A modern graphics calculator can give all solutions quickly (and thus efficiently), as the two screen dumps from a Casio cfx-9850G calculator below suggest:

\[
\begin{align*}
\text{ax}^3 + bx^2 + cx + d &= 0 \\
0.9902 - 1.1917i
\end{align*}
\]
Analytic and graphical solutions of such an equation are slow and less comprehensive, and thus much less efficient. In contrast, the calculator provides both real and complex solutions, almost immediately after the coefficients are entered.

To illustrate situations for which graphics calculators might be expected to be used by some students and not by others, consider the task of investigating the effect of an annual compound interest rate of 6%, starting with an initial amount of $14,000. For the related task, a student may choose to use their calculator as shown below:

\[ \text{Ans} \times 1.06 \]
\[ 14000.00 \]
\[ 14848.00 \]
\[ 15730.48 \]
\[ 16641.22 \]
\[ 17574.66 \]
\[ 18735.16 \]

Another student may prefer to use a suitable compound interest formula (involving an exponential function). The choice may depend on the level of comfort a student enjoys both with the calculator and with the relevant mathematics. It may even be that somebody quite comfortable with both the relevant mathematics and with the calculator may choose one of these on some occasions and the other at other times. For instance, the calculator solution would be tedious and error-prone if the amount required involved a long time (say twenty years) into the future. On the other hand, a formula here would provide information about the final value but no information about annual changes.

There are several circumstances in which we would expect calculators not to be used in assessment situations:

- exact answers are required;
- symbolic answers are required;
- written explanations of reasoning are required;
- task involves extracting the mathematics from a situation or representing a situation mathematically;
- graphics calculator use is inefficient;
- task requires that a representation of a graphics calculator screen will be interpreted.

To illustrate one of these cases, students might be asked to interpret the Texas Instruments TI-82 calculator screen below, to find a value of \( x \) for which \( f(x) \) has an approximate value of 6.

\[
\begin{array}{c|c}
 x & y1 \\
\hline
 0 & 1.1052 \\
 1 & 2.4026 \\
 2 & 5.0673 \\
 3 & 8.2435 \\
 4 & 10.535 \\
\end{array}
\]

\[ y1 = e^{0.1x} \]
It is worth noting that there are sometimes difficulties in classifying items with such a typology. For example, a task may be written to apparently demand that analytic methods be used (in preference to a calculator), when in fact a competent student may use their graphics calculator to thwart this intention. For example, consider the task:

Solve exactly \( e^{x^2-5} = 1 \)

The use of the word 'exactly' suggests that a numerical method is inappropriate. In this case, the numerical result of \( \sqrt{5} \) is obtained on a TI-82, using a \texttt{solve} command, and an astute student can square this result to see that it must be the square root of 5:

\[
\texttt{solve}(e^{x^2-5}-1, x, 0) = 2.236067977
\]

The construction of examination tasks is clearly important, both to communicate intended student behaviours and also to ensure that the performance expected is being prompted. To illustrate this point, consider the following theme and variations on an examination task discussed by Kissane, Kemp & Bradley (1996):

Solve \((x + 2)^2 = 9\)

Solve \((x + 2)^2 = 8\)

Solve exactly \((x + 2)^2 = 8\)

Solve \((x + a)^2 = 9\)

Solve \((x + 2)^2 = 9\) and explain why there are only two solutions.

These examples suggest that tests and examinations in which students have graphics calculators need to be written carefully — as indeed do all formal assessment tasks, of course. The relationship between a test and the use of graphics calculators ought be a conscious \textit{design}, not just occur accidentally, as suggested by Kemp, Kissane & Bradley (1996). As well as helping analyse and understand student responses to examination tasks, this typology may reveal our intentions or habits. Bradley (1999) provides several examples of graphics calculator use in a recent high stakes examination.

\textbf{What is to be recorded?}

Another issue concerns the fact that it is usually hard for students to write down adequate steps to 'show their working' on questions for which they have made significant use of a graphics calculator. Usually, a listing of calculator keystrokes is not particularly helpful (although it can be quite illuminating provided the person reading it is thoroughly familiar with the particular calculator used.)

Kissane, Bradley and Kemp (1994) have suggested two general principles for this issue and discussed them briefly:
If we require working to be shown, it should be worth showing in its own right, and not only as a means of awarding partial credit.

If only part marks are to be awarded for numerically correct answers, for which working is not provided, then this should be stated explicitly in advance.

**Equity in assessment**

Discussions about the place of graphics calculators in assessment rarely progress far without some mention of issues of 'equity'. In some states, the widespread (and official) unease about 'equity' issues is largely responsible for a delayed introduction of graphics calculators into schools. In other states, a concern for 'equity' has strongly influenced assessment practice, especially in high stakes external examinations. Several issues associated with equity have been canvassed in Kissane, Bradley and Kemp (1994).

A key issue is whether or not it is equitable to **not** allow graphics calculators in external examinations. As Kissane, Bradley & Kemp (1996) argue, one consequence of preventing students from using graphics calculators in examinations is that students in the least-resourced schools are consequently denied access to this useful technology. There is rarely money for 'frills' in such schools, but resources will be found for 'necessities'. When an examination is prepared on the assumption that students have access to graphics calculators, and the corresponding course has been designed accordingly, decisions can be made to allocate funds suitably at the school level. It is much harder to argue for appropriate resourcing at the school level for a technology that is merely 'recommended'. Around Australia, school budgets frequently contain substantial provisions for Information Technology resources, usually in the form of computer laboratories and associated telecommunications networks. A small shift of some of these funds towards the much less expensive technology of graphics calculators is possible, when graphics calculators are seen as necessities rather than as luxuries. These issues are clearly less of a concern for schools that are well resourced, and are well located within relatively affluent communities.

There are many different models of calculators, and so attention is needed to ensure that students are not unduly advantaged or disadvantaged for assessment purposes with respect to their peers by their particular choice of calculator (which may not have been their own choice, of course). Prudence suggests that both schools and individuals may need some careful guidance to ensure that underpowered calculators are not purchased unknowingly and also that students are not unduly advantaged over others because of especially powerful calculators (which usually cost more to purchase, of course). Suitable specifications are needed to achieve these two ends, as described briefly in the next section.

All graphics calculators are programmable, and it seems sensible to allow the programmability to be used to upgrade calculator capabilities where they are comparatively modest. As noted by Kissane, Bradley & Kemp (1994), this strategy was originally used in the Advanced Placement examinations of the College Board in the USA, where minimum calculator specifications were demanded of calculators, and programs used to add these in the case of less sophisticated calculators. Of course, such a strategy is of no value if calculator memories are required to be cleared before an examination commences, which is why such a practice is most undesirable.
Calculator memories that allow programs to be added, thus potentially reducing some differences between models, also allow text to be entered, in some cases in an awkward way. One response to this possibility is to demand that memories be cleared before examinations begin, but this seems to eliminate any possibility of calculator upgrading through the use of programs. An alternative strategy, employed in some places to apparently good effect, is to allow students to take some pages of notes with them into an examination. Such a practice raises some interesting issues regarding the nature of mathematical activity; for example, how critical is human memory in mathematics?

The use of 'calculator-neutral' testing to protect equity seems particularly problematic. In the first place, it is not clear that it is actually possible to construct (valid) calculator neutral tests. In the second place, the use of such an approach sends an odd signal to the community: that graphics calculators are 'useful' enough to encourage their use and purpose, but not useful enough to be of any advantage in an examination. It is very difficult to make such a message convincing.

It needs to be acknowledged that students who have regular and personal access to a graphics calculator (e.g. one that they personally own or that is theirs on long term hire or loan from a school) would seem more likely to make good use of it and to become familiar enough with it and use it well, than are those students whose physical access is restricted to the use of a class set of calculators handed out by the decision of the teacher and collected up again at the end of a lesson.

In recent years, it has become evident that equity issues are not restricted to access to the physical technology, however, but also concern access to quality guidance and support. There are at least three aspects of this: students need to have teachers who are competent users of graphics calculators themselves, so that they may get suitable instruction in their appropriate use; they need school mathematics courses to be taught in a way that integrates technology into them, rather than regarding it as an occasional add-on; they need curriculum materials which integrate calculators into the mathematics. These three aspects of equity are quite difficult to deal with. None is easy to bring about, and none is evident at first, since attention often is focussed on the provision of hardware.

**Calculator specifications**

As noted in the previous section, for equity reasons, the range of allowable calculators in formal assessment settings is often restricted. A common example of this is that calculator keyboards with a QWERTY keyboard are banned from use in examinations (although of course that does not mean that they could not be used at a local level within a classroom or even within an assessment context where all students had comparable access to them). The banning of calculators with QWERTY keyboards (originally by the Advanced Placement examination authorities in the USA) clearly reflected a concern for easy text entering and recall. However, such an indirect specification seems to be a stop-gap measure, since it does not address precisely the concern at hand: the power of the technology and the possibility of students having differential access to it. Indeed, it now seems a little ironic that this specification, which prevents the Texas Instruments TI-92 calculator from being used in the relevant examination, does not prevent the use of the companion TI-89 calculator, which is more powerful in terms of capabilities — but lacking a QWERTY keyboard. (Readers interested in regulations regarding calculator specifications will find some links to them on the web site: http://www.staff.murdoch.edu.au/~kissane by
following the links to graphics calculators.)

A more prudent approach to dealing with calculator specifications may be more direct, such as not permitting calculators with 'Extensive' symbolic manipulation' to be used in high stakes examinations. While this still requires careful definition of both 'extensive' and 'symbolic', usually by reference to particular calculator models, it seems more sensible than hoping other attributes (such as a QWERTY keyboard) will carry the information suitably. Of course, all calculators provide a form of symbolic manipulation; after all the numerals are symbols for their corresponding numbers. Assessment regulations that bar any symbolic manipulation are unlikely to help.

An emerging issue for calculator specifications concerns 'flash memory', now available on a number of calculators. The capabilities of calculators with flash memory can be upgraded as new features become available (or are seen to be popular, or even desirable). While this appears on the one hand to be a good strategy for calculator design, it creates a new problem for high stakes assessment, since the calculator capabilities cannot be determined without actually operating it. While this is also true to an extent of any programmable calculator, at least it is fairly well understood what kinds of programs are available (and perhaps more importantly, what kinds are not available). A decision on suitability can be made by examination authorities. Typically, programs for calculators do not add significantly to the basic inbuilt capabilities of a calculator, but rather make use of them in tailored ways. It is much harder to deal with this issue when it is the capabilities themselves that are being upgraded. As for programming, flash memory upgrades are not restricted to those provided by calculator manufacturers, so that reliable informed advice on the possibilities is especially difficult to obtain.

The last few years have seen the development of algebraic calculators, which have extensive symbolic manipulation capabilities built in to them, of particular relevance to algebra and the calculus (see Klessie (1996; 1999) for some examples of these). The screen below shows a few algebra examples from the Texas Instruments TI-92.
The next diagram shows a succession of screens generated by the Casio Algebra fx 2.0.

The assessment issue related to such calculators is that it is much harder to distinguish student thinking from machine work and thus to interpret what is written down in a formal assessment setting. Algebraic calculators are significantly more expensive (today) than other kinds of graphics calculators, and so there are more likely to be equity issues associated with their use. Without doubt, dealing with the assessment issues associated with algebraic calculators is one of the biggest challenges facing curriculum and assessment authorities in the near future. Such problems are less acute at the local level (e.g. at a single school), where it is much easier to observe, control and understand the access and equity issues.

**Electronic algorithms**

Another assessment issue is that some calculator programs seem to be designed mainly to allow exam questions to be answered. These are often fairly large programs, but do not require students to key them in, since electronic transfer of programs is possible with most calculator models these days. The screens below show a calculator program of this kind for handling routine questions concerned with geometric progressions.

Such programs are conceptually similar to routine algorithms, requiring only that students recognise key aspects and perform a routine computational procedure to obtain the desired answer. It does not seem as if much useful assessment information can be gained from student responses to routine questions that can be dealt with in this kind of way. (Of course, it is arguable that the same is true of paper and pencil algorithms of the same kind.)
The most obvious way of dealing with such possibilities is to not ask routine questions of a kind that can be answered numerically like this. Another is to demand that students 'explain' their working in some sense (although this is not always easy to ask — or to do). It is important to note too that there is a substantial intellectual difference between students writing such programs for themselves (which will usually reflect significant learning) and using programs written by others (which often will not reflect much learning of mathematics at all).

Assessing calculator usage

Mathematics assessment is concerned with how well students know, understand and are able to do mathematics. But in a context in which technology is important, we may also wish to assess how well students can use a graphics calculator (without misunderstanding this as mathematics). It is important for students to make efficient use of technology, although it is probably easier to assess this informally than formally. As already noted, it is quite hard for students to 'show their work' when using a calculator; however, it is relatively easier to watch students using calculators and appraise the efficiency and sensibleness with which they do so. Assessment of calculator use seems to be important, if teachers are to provide useful feedback to students to improve their performance in this respect.

In summary

Although many of the issues raised here involve formal testing, assessment ought to involve (much) more than testing. Some coherence between assessment practices and the mathematics curriculum is critical. Graphics calculator use in assessment is desirable and even necessary in the quest for this coherence, but will demand some rethinking of assessment practice. Although it has always been important for tests to be carefully designed, it seems that the possibility of calculator use in tests makes this even more critical.

Finally, two observations about assessment, made almost a generation apart, offer some food for thought when considering the many assessment issues related to graphics calculators. The first serves to remind us of the considerable limitations of formal assessment:

Contemporary assessment recognises the inadequacy of the 'Assessment as measurement' metaphor. Our goal is now 'Assessment as portrayal' (Clarke, 1997, p. 65).

The second reminds us of the significance of assessment in the context of teaching and learning mathematics:

Nobody ever got taller just by being measured (Cockcroft Report, 1980).
References


Note that several of the above publications are available from the following web site: http://www.staff.murdoch.edu.au/~kissane/epublications.htm

Acknowledgement

I acknowledge the contributions to the thinking in this paper of both Marian Kemp and Jen Bradley of Murdoch University. The three of us have worked on aspects of assessment with graphics calculators over several years.