Morgan, C. (2000). Evaluation of pupils' mathematical learning – a critical view. In E. Fernandes & J. F. Matos (Eds.), *Actas do ProfMat 2000* (pp. 134-144). Funchal, Madeira: Associação de Professores de Matemática.

Evaluation of pupils' mathematical learning – a critical view

Candia Morgan

Institute of Education, University of London, United Kingdom

In July this year, I had the interesting experience of visiting Japan. Apart from what I gained by attending a conference of mathematics education researchers, one of the most lasting impressions of the few days of my visit was the feeling of disorientation and incompetence aroused by my inability to read the everyday texts (such as road signs, menus, train timetables) surrounding me, to speak and to practise or even recognise appropriate forms of behaviour. Even when conversations took place in what appeared to be a common language, there were occasions when it seemed that similar forms of words meant entirely different things to participants from different cultures and that questions and answers held little relationship to one another. Of course, cross-cultural difficulties in communication are not restricted to such exotic locations. Indeed, I intend to argue that they also occur in apparently monolingual and monocultural settings.

You may ask: What has this to do with evaluation of mathematical learning? I wish to focus here on the evaluation by their teachers of the mathematical learning of students in school, university or other formal educational settings. Any such evaluation depends fundamentally on an act of communication. Students may be aware that evaluation is taking place and may consciously attempt to shape their text production (in written, spoken or other form) to present a positive impression. On the other hand, texts that students produce with other intentions in mind may also be subject to evaluative interpretation by their teachers. Indeed, teachers may evaluate their students' learning without being aware that they are doing so, forming 'impressions' that are never recorded or even recognised as evaluations, yet serve to influence future interpretations of the students' behaviours. In any of these cases, the texts produced by a student are necessarily shaped and constrained by the resources that the student brings to the task.

By resources, I mean not only the student's mathematical knowledge but also their linguistic and cultural knowledge, their personal history of experience in and out of school, their beliefs about the nature of the activity in which they are engaged and about the types of text and behaviour that are appropriate in this context. Similarly, the teacher's interpretation and evaluation of each text is shaped by the resources they bring.

When we, as teachers, evaluate our students, our evaluations are generally underpinned by an assumption that we are members of the same culture, speaking the same language. I would suggest that this assumption is usually made even in obviously multi-cultural and multi-lingual classrooms. This means that we tend to be confident in the judgements we make of students' meanings and intentions or, where we find a text unclear, we lay the blame on the inadequate expression of an inexperienced or 'weak' student. We do not often imagine that we and our students might be natives of different countries, speaking different languages, using and valuing different conventions of behaviour.

Mathematics education discourse increasingly emphasises the valuable formative role that assessment can play. According to the National Council of Teachers of Mathematics (NCTM) in the United States, "[supporting] the continued mathematical learning of *each* student … is the central goal of assessment in school mathematics" (1995, p. 6, original italics). Formative assessment may also play other positive roles:

> The purpose of the assessment process is to make explicit children's achievements, celebrate their achievements with them, then help them to move forward to the next goal. … When shared with the child, assessment information is more likely to result in a raising of standards, because the child is more focused, motivated and aware of his or her own capabilities and potential. Good assessment practice enables children to be able to fulfil their learning potential and raises self esteem and self-confidence. (Clarke & Atkinson, 1996, p.9)

Indeed, some types of formative assessment do appear to have measurable effects on performance, the quality of the feedback that is provided to students being a key factor (Black & Wiliam, 1998).

Morgan, C. (2000). Evaluation of pupils' mathematical learning – a critical view. In E. Fernandes & J. F. Matos (Eds.), *Actas do ProfMat 2000* (pp. 134-144). Funchal, Madeira: Associação de Professores de Matemática.

It seems common sense that the more aware both teachers and students are of what learning has taken place the better placed they will be to develop future learning. This celebration of assessment, however, ignores its other features and consequences. I have argued elsewhere that it is important to be aware (and to beware) of the regulative and discriminatory functions that all kinds of evaluation of mathematical achievement play in society (Morgan, 2000a, 2000b). What I want to point to here specifically is the potential for failure in the communicative/evaluative process and the issues that this raises for mathematics teachers and learners.

Examples of evaluative readings – and alternative interpretations

It is easy to find examples in the mathematics education literature of evaluative readings of students' spoken or written mathematics. It is also often acknowledged that there are alternative ways of interpreting what students have said or done.

My first example is taken from a review of research by Dickson, Brown and Gibson (1984, p. 337) in which they discuss the types of errors that students can make. Beneath a picture of twelve children, the student is asked:

> I have 24 lollies and I want each child to have the same amount of lollies. How many lollies will I give each child?

I am confident that all of us here can recognise what mathematics is required in order to answer this question correctly and perform the required calculation accurately. Dickson et al. report that Percy, a 12-year-old, answered 'one' because:

> I would give each child one lolly and keep 12 for myself.

One approach to evaluating this response is to attempt to diagnose the point in the solution of the problem at which Percy's competence breaks down. It is well known from teachers' everyday experience and from research that 'word problems', in which a task involving one or more computational step is presented in the form of a 'story',

present more difficulties to students than the corresponding calculations would if they were presented without a context. One approach to explaining these difficulties conceptualises the task as one of 'decoding' the words into an appropriate symbolic form or of 'extracting' the required numbers and operations from the verbal text. Achieving a solution is seen as a determinate linear process: read, extract the information (numbers, operation) needed from the text, form the appropriate equation, perform the required calculation. An interpretation of the cause of a student's failure to produce the approved answers is formed by observing at what point this process breaks down. However, such separation of mathematical structure from the context in which the problem is presented appears over simplistic (particularly in the light of research such as Cooper & Dunne, 2000). In Percy's case, the problem cannot simply be seen to arise at the point of extracting the required operation from the text. In one sense he has done so correctly.

Dickson et al. suggest that Percy's answer should be considered correct, because the form of the question is ambiguous. In one way, I would agree with this evaluation. In another, however, it is naïve because it takes no account of the culture of most mathematics classrooms for most 12-year-olds. 'Insiders' in that culture (who include most children and their teachers) know that Percy is wrong, regardless of the inadequate wording of the question. While he has succeeded in distributing the sweets, he has not succeeded in performing the task that was expected of him. Did Percy answer in this way because he did not know the conventions of this type of question or because he was deliberately refusing to play the game? Does he know that 24 divided by 12 would be an appropriate model for this problem? Does he know that 24 ÷ 12 equals 2? Cockburn, commenting on this example, draws attention to the need to be careful about the assumptions we make about students' understanding on the basis of such responses, suggesting that Percy "did not possess, or chose not to use, experience in 'playing the

game'" (1999, p. 9). I would argue that it is not simply a question of experience – after all, a 12-year-old must have had plenty of experience of the game of school mathematics, even if he usually loses. It must also involve a grasp of the rules of the game, the 'ground rules' of school mathematics practices (Edwards & Mercer, 1987; Sheeran & Barnes, 1991). By this stage in his school career, Percy probably knows very well that he is a foreigner in the mathematics classroom – recognising that he cannot get the approved answers/speak the language, but unable or unwilling to find a means of doing so.

In this example, it is fairly clear that the answer the student has produced is not that required by the mathematical structure of the task. Mathematics teachers (who have, after all, been successfully inducted into the rules for playing the game of school mathematics) will reliably identify that the answer is not legitimate. Alternative evaluations of the student's mathematical thinking are based on different suppositions about the source of the 'error'. It is likely in Percy's case, however, that, irrespective of his teacher's particular interpretation of the root cause of his difficulty, similar (probably appropriate) teaching would be provided, including attention to the knowledge and skills needed to recognise the conventions of 'word problems'. In cases where the set task is more open, identification of the legitimacy of student responses may be more problematic and the range of possible evaluations of student thinking may be accordingly more diverse. Moreover, the consequences of such different evaluations may be significantly different.

I will give two examples from my study of teachers assessing students' written reports of their work on investigative tasks (Morgan, 1998). The students had worked on a task called "Topples" that involved investigating piles built up of rods of increasing lengths, seeking a relationship between the length of the rod at the bottom of the pile and the length of the rod that would first make the pile topple over. One girl, Sandra, had drawn

numerous diagrams, showing all the piles she had experimented with (see Figure 1). These piles were beautifully drawn: neat and accurate in all details, including the colours of the Cuisenaire rods she had used, and indicating where she had had to use two rods to make up a given length because longer rods were not available.

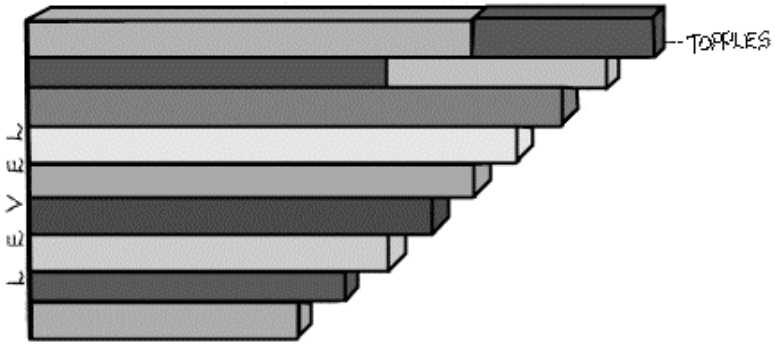

Figure 1: One of Sandra's piles of rods

Several of the teachers who evaluated Sandra's work commented positively on the aesthetic qualities – but in the same breath stated that she had wasted her time on unnecessary detail (something that was identified by some as typical of girls). The diagrams also seemed to prompt further low evaluations of Sandra's way of working and even her general intellectual capacity. Although all the students involved must have used 'split' rods in order to generate the results they reported, only Sandra showed them in her diagrams. This led some teachers to the conclusion that she was working at a low level, in a very practical way. The possible consequences for Sandra's future experiences in her mathematics lessons are indicated by the diagnosis and remedy suggested by one teacher:

> I would … get some blocks of wood and chop them up and see … if it does make a difference. . I would get actually longer lengths um to try and eradicate this. So that's what's done it I think. That's what's caused the damage.

The practical aspects of other students' work did not give rise to comment because they were hidden by the use of fewer, more abstract diagrams. If Sandra had not used such realistic diagrams, it is possible that her work would have been evaluated differently,

that different meanings would have been attached to errors in her results and different future learning opportunities provided for her. We may speculate about her reasons for drawing these unfortunately good diagrams. Doing lots of repetitive examples and presenting them neatly and accurately are highly valued in many school mathematics contexts. An alternative reading of Sandra's text might suggest that she simply does not recognise the different rules of the game in the context of more open work where higher level mathematical thinking and creativity are more highly valued. What difference would it make to our actions as teachers if this were our interpretation of the situation?

The source of different interpretations cannot always be located in the student's grasp of the rules of the particular mathematical discourse. Teachers' expectations about mathematical content may also affect the ways they bring their resources to bear on their evaluation of a student text. Another student Steven worked on the same "Topples" task. Having found a formula $(A + A) + \left(\frac{A}{2}\right) = b$ expressing the required relationship (accepted as legitimate by all the teachers who read it) and shown that he could use it to find the length of the 'topple rod' for piles starting with rods longer than those he had available to build with, Steven presented an alternative method for piles starting with long rods by scaling up his results for piles starting with short rods (Figure 2).

$$(100 + 100) = 200 \quad \left(\frac{100}{2}\right) = 50$$

$$200 + 50 = 250$$

250 would be the one at which it would topple.

An alternative way to do this would be to take the result of a pile starting at 10 and multiply it by 10.

$$(10 + 10) = 20 \quad \left(\frac{10}{2}\right) = 5$$

e.g. $20 + 5 = 25$

$$25 \times 10 = 250$$

Morgan, C. (2000). Evaluation of pupils' mathematical learning – a critical view. In E. Fernandes & J. F. Matos (Eds.), *Actas do ProfMat 2000* (pp. 134-144). Funchal, Madeira: Associação de Professores de Matemática.

or you could even take the basic result for 1 without rounding it up and you could multiply it by 100

$$(1+1) = 2 \quad \left(\frac{1}{2}\right) = 0.5$$

e.g. $2 + 0.5 = 2.5$

$$2.5 \times 100 = 250$$

Figure 2: Steven's application of his original formula and an alternative method

No further justification of this method was provided in the text. In particular, there was no indication of how it had been derived. In order to evaluate Steven's work, each teacher-reader constructed his or her own understanding, not only of the method itself but also of the means by which Steven might have derived it. The range of possible evaluations is illustrated by extracts from interviews with three teachers reading this section of Steven's text.

> Teacher 1: Charles. Um ok, so I mean he's found the rule and he's quite successfully used it from what I can see to make predictions about what's going to happen for things that he obviously can't set up. So that shows that he understands the formula which he's come up with quite well, I think. There's also found some sort of linearity in the results whereby he can just multiply up numbers. Which again shows quite a good understanding of the problem I think.

Charles legitimates the alternative method mathematically, relating it to the linearity of the relationship between the variables. He takes this as a sign that the student has "come up with" the formula as a result of understanding the linearity of the situation. This results in a positive evaluation of Steven's mathematical understanding.

> Teacher 2: Grant. It's interesting that the next part works. I don't know if it works for everything or it just works for this, but he's spotted it and again he hasn't really looked into it any further. He's done it for one case but whether it would work for any other case is, er, I don't know. He hasn't looked into it … And he's used it in the next part, er, used the this multiplying section in the next part, and it's just a knowledge of number that's got him there, I think, intuition whatever. He may have guessed at a few and found one that works for it.

Grant appears less confident with the mathematical validity of the "alternative" formula, expressing uncertainty about whether the method would work in general. Perhaps

because of this uncertainty, his narrative explaining how Steven might have arrived at the method devalues the student's achievement, suggesting that the processes involved were not really mathematical: spotting the method, not looking into it properly, guessing, using "just a knowledge of number" or intuition. A negative evaluation is given both of Steven's result and of the processes he may have gone through in order to arrive at it.

> Teacher 3: Harry. And he's got another formula here …. I don't really understand what he's done here…. So he's produced another formula where …. He's taken the result of a pile starting at ten and multiplying by ten and I don't understand what he's done there…. I would have asked him to explain a bit further. He's – the initial formula with two hundred and fifty is proved to be correct and he's trying to extend it. He's trying to look for other ways. Maybe he has realised that two hundred and fifty could be the exact answer or maybe not. So he's trying other ways to explain some of the inconsistencies that he's seen, but I think greater explanation needed here.

Like Grant, Harry seems to have some difficulty making sense of the mathematics. In spite of this, he is able to compose yet another narrative to explain the student's intentions, stressing by repetition the suggestion that Steven has been "trying" (possibly with the implication that he has not succeeded). Harry, although appearing willing to approve Steven's perseverance, also has a low evaluation of his mathematical understanding, even questioning whether he knows that his first method had already yielded a correct answer.

In order to make sense of a text and to use it to evaluate the student-writer's achievement, each teacher must compose an explanatory narrative, drawing on the resources available to him or her. These resources include common expectations of the general nature of mathematical problem solving and reports of such work as well as more personal mathematical understanding and experiences. The three teachers' different interpretations of Steven's level of understanding and different hypotheses about the methods he might have used are related to their personal mathematical resources. Charles legitimates the mathematics Steven has used, identifying it with the

high level mathematical concept of linearity. It is he who makes the most positive evaluation of Steven's understanding and his success in solving the problem. Grant and Harry, on the other hand, express uncertainty about the general validity of the formula. Both these teachers construct pictures of the student working in relatively unstructured or experimental ways – a much lower evaluation of his mathematical thinking. While I would argue that it is possible to say that Steven's alternative method is mathematically equivalent to his original formula (a point not recognised by Grant or Harry), this alone does not make it possible to say how Steven arrived at his conclusions or how he himself understood them.

It is possible that later discussion with Steven would have helped each of these teachers to clarify the nature of his conclusion and of the mathematical thinking that led to it, possibly changing their evaluations – if time was available for such in depth discussion. Without this, however, the different evaluations would be likely to lead not only to different grades for this piece of work (and possibly high-stakes consequences arising from this) but also to different expectations of Steven's achievement and consequently different treatment of him in the classroom.

Discussion

Rather than assuming that communication between student and teacher is transparent, it would seem that alternative interpretations of student texts are always possible. So what is the teacher to do if she is to evaluate students' achievement in ways that are fair, useful and practical? In my opinion, one prerequisite is to address the rules of the game explicitly with students in the classroom. Learning mathematics is not just about learning concepts, facts and skills; it also includes learning about mathematical ways of thinking and communicating, learning how to produce texts that will be judged to be mathematical.

Of course, even when students know the rules of the game there is no guarantee that they will play according to the rules. In some cases, they will openly refuse to co-operate. This is relatively easy for teachers to identify – if not so easy to deal with! Opting out of the game makes it clear that, while the student's behaviour may be evaluated according to the social norms of schooling, the futility of any attempts to evaluate mathematical learning is obvious to all concerned. Alternatively and, I suspect, more frequently, students, knowing that they will be judged by the results achieved rather than by the means of achieving them, will strive to produce texts that will be recognised as legitimate texts by their teachers but that are produced in ways that would not be recognised as legitimate, for example, by copying answers, reconstructing legitimate methods from answers achieved in non-approved ways, or faking data to fit known results. Students may do this with the purpose of making space and time for themselves to pursue their own goals rather than those intended by the teacher. Tomazos' (1997) study of the students' perspective in one primary classroom describes students who consciously, deliberately and successfully concealed their disengagement from the mathematical goals of the tasks set by their teacher. As Tomazos comments, "the remarkable feature of this is that the process can be invisible to teachers and researchers alike" (p. 228). Whether you choose to call it cheating, dissembling or resistance it has serious consequences for the meaningfulness of any form of evaluation of mathematical achievement.

Successful students learn to recognise the characteristics of high status mathematical texts and are able to produce texts of their own that will lead to high evaluations of their mathematical achievement. Other students do not. As in other aspects of schooling, it is likely that students from relatively privileged backgrounds, sharing the culture and language of the school, are over-represented among those who succeed. Many from non-hegemonic groups are disadvantaged because they lack the cultural and linguistic capital

(Bourdieu, 1998) that would enable them to gain recognition. Although a classroom may appear to be monolingual and monocultural, not all students are natives to the culture of the mathematics classroom. Evaluating their achievement without recognising this is like judging the behaviour and speech of the newly-arrived foreigner in your midst without recognising their ignorance of your language and culture.

References

Black, P. & Wiliam, D. (1998). *Inside the Black Box: Raising standards through classroom assessment*. London: King's College, School of Education.

Bourdieu, P. (1998). *Practical Reason*. Cambridge: Polity Press.

Clarke, S. & Atkinson, S. (1996). *Tracking Significant Achievement in Primary Mathematics*. London: Hodder & Stoughton.

Cockburn, A. D. (1999). *Teaching Mathematics with Insight: The identification, diagnosis and remediation of young children's mathematical errors*. London: Falmer Press.

Cooper, B. & Dunne, M. (2000). *Assessing Children's Mathematical Knowledge: Social Class, Sex and Problem-Solving*. Buckingham: Open University Press.

Dickson, L., Brown, M., & Gibson, O. (1984). *Children Learning Mathematics: A Teahcer's Guide to Recent Research*. Eastbourne: Holt, Rinehart and Winston.

Edwards, D. & Mercer, N. (1987). *Common Knowledge: The Development of Understanding in the Classroom*. London: Methuen.

Morgan, C. (1998). *Writing Mathematically: The Discourse of Investigation*. London: Falmer.

Morgan, C. (2000a). Better assessment in mathematics education? A social perspective. In J. Boaler (Ed.), *Multiple Perspectives on Mathematics Teaching and Learning* (pp. 225-242). Westport, CT: Ablex Publishing.

Morgan, C. (2000b). Discourses of Assessment – Discourses of Mathematics. In J. F. Matos & M. Santos (Eds.), *Proceedings of the Second International Mathematics Education and Society Conference* (pp. 58-76). Lisbon: Centro de Investigação en Educação, Faculdade de Ciências da Universidade de Lisboa.

NCTM. (1995). *Assessment Standards for School Mathematics*. Reston VA: National Council of Teachers of Mathematics.

Sheeran, Y. & Barnes, D. (1991). *School Writing: Discovering the Ground Rules*. Buckingham: Open University Press.

Tomazos, D. (1997). Investigating change in a primary classroom: valuing the students' perspective. In E. Pehkohnen (Ed.), *Proceedings of the 21st Conference of the*

Morgan, C. (2000). Evaluation of pupils' mathematical learning – a critical view. In E. Fernandes & J. F. Matos (Eds.), *Actas do ProfMat 2000* (pp. 134-144). Funchal, Madeira: Associação de Professores de Matemática.

International Group for the Psychology of Mathematics Education (Vol. 4, pp. 222-229). Lahti: University of Helsinki, Lahti Research and Training Centre.