Learning to prove: using structured templates for multi-step calculations as an introduction to local deduction

Tony Gardiner, Birmingham (Great Britain)

Abstract: It is generally accepted that proof is central to mathematics. There is less agreement about how proof should be introduced at school level. We propose an approach - based on the systematic exploitation of structured calculation - which builds the notion of objective mathematical proof into the curriculum for all pupils from the earliest years. To underline the urgent need for such a change we analyse the current situation in England - including explicit evidence of the extent to which current instruction fails even the best students.

Kurzreferat: Es ist allgemein akzeptiert, dass Beweisen ein zentrales Thema der Mathematik ist. Weniger klar ist, wie dieses Thema im Rahmen des Unterrichts in der Schule behandelt werden sollte. In diesem Beitrag wird ein Weg vorgeschlagen, der auf der systematischen Nutzung eines strukturierten Kalküls basiert. Damit ist die Idee des mathematischen Beweises im Unterricht bereits zu einem sehr frühen Zeitpunkt in den Unterricht zu integrieren. Um die Bedeutung des Themas für den Unterricht zu belegen, wird die derzeitige Situation an Schulen in England betrachtet, die als wenig hilfreich selbst für manche gute Schülerinnen und Schüler angesehen wird.

ZDM-Classifikation: C30, C50, C70, D30, D50, E50, N70

1 Identifying the problem

Many national systems include explicit statements about the importance of "proof" in school mathematics. But some appear confused about how to reflect the acknowledged importance of this topic within the curriculum - in what is taught, in what is expected of students, and in how material is assessed.

English 18 year olds have never found it easy to construct proofs. But university mathematicians in the UK have in recent years observed a marked change in school-leavers who choose to study highly numerate disciplines: "Most students entering higher education no longer understand that mathematics is a precise discipline in which exact, reliable calculation, logical exposition and proof play essential roles; yet it is these features which make mathematics important" (LMS 1995, p. 8); "What is being observed in England *is* a profound change. There may be those in certain other countries for whom such behaviour is all too familiar. That is no reason to dismiss the fact that England is now witnessing mathematical behaviour of a kind never previously experienced in the top 10-20% of the ability range" (Gardiner 1995, p. 343). The very conception that mathematical calculations have to be structured logically if they are to be trusted has been largely replaced by a blind faith in half-remembered "rules" - often used in a form which appears to be made up on the spot (Gardiner 2003a, 2003b). The central problem now is not that students have difficulty *constructing* proofs, but rather that they have no conception that the essence of mathematics lies in *exact calculation* and proof, what constitutes a proof, what distinguishes exact calculation from apparently intelligent inference, and why these things matter.

The confused situation in England with regard to proof would appear to reflect a wider failure within the mathematics education community to understand the nature of proof. Mathematics is an open discipline, with a public procedure (of which "proof" is a key component), which requires new developments to be published in a form that can be scrutinised - and hence validated, refined or rejected - by Everyman. This spirit of openness includes an on-going debate about, and a permanent re-evaluation of, the procedure itself. Unfortunately, this openness is exploited in the chapter (Hanna 1996) on "Proof and proving" in the *International handbook of mathematics education* to give the impression that the difficulty of developing an effective approach to proof *within classroom practice* arises "because there have been and remain differing and constantly developing views [within mathematics itself] on the nature and role of proof and on the norms to which it should adhere" (p. 877). Nothing could be further from the truth. The challenge of finding an effective way to incorporate proof as a key component in school mathematics is only obscured by such attempts to shift the focus of attention away from the real source of the present confusion, which lies squarely within mathematics education.

The confusion within mathematics education is exacerbated by blurring the distinction between two very different notions - namely "logical correctness" and "psychological conviction". This has led to thoroughly misleading claims, such as "a proof actually becomes legitimate and convincing to a mathematician only when it leads to real mathematical understanding" (Hanna 1996). A proof may well provide a degree of insight and understanding. But the official role of proof in mathematics is *purely* to demonstrate logical correctness. While there is scope for exploring how this *logical* function of proof is best incorporated in school mathematics, analyses based on a blurring of the distinction between objective correctness and subjective conviction have contributed nothing but unhelpful confusion in the minds of both teachers and students.

1.1 The background

The observed change in English students' understanding of proof is difficult to document *post hoc*. Yet it is sufficiently clear to those who work with university mathematics students to oblige one to look for possible causes - even though such a profound change is unlikely to have a single cause. Thus the attempt in this section to identify the origins of our current dilemma is likely to be partial and to some extent speculative; and any conclusions are inevitably tentative. We therefore beg the reader's temporary indulgence.

Once the question of possible causes is raised, there is no escaping from the fact that the reported change followed hard on the heels of the spread of dependence on calculators in the 1980s and 1990s, and the introduction in 1988/9 of the first ever English national curriculum and of the unified GCSE system of public examinations at age 16.

The impact of calculators has been profound but elusive. In particular, their influence is not restricted to the way they are actually *used*, but rather lies in the attitudes that their style of use engenders; this becomes clear only when one encourages students to talk through their attempted written "solution procedures" (for example, in connection with the three problems discussed in section 1.4). A possible summary of what one observes among 18 year old university entrants is that the majority have been allowed to depend on the power of the calculator from too early a stage, with the result that they have been effectively encouraged to think of "solving a mathematical problem" in terms of adopting an almost random sequence of hopeful attempts aimed solely at "getting an answer". The fact that calculators allow many students to achieve apparent success in this manner has consequences for the whole of school mathematics. The most significant consequence is that such use of the calculator now routinely confers "success" without requiring students to achieve conceptual understanding, and so inculcates a belief - in students and teachers - that it is sufficient for the learner to operate on the crudest imaginable conceptual level: this is well illustrated by the student responses to the first and second problems in section 1.4. A more elusive consequence would appear to be the loss of any imperative for the student to achieve, or for the teacher to demand mastery of even the simplest universal procedures: this is well illustrated in the third example in section 1.4 and in (Gardiner 2003a, 2003b).

The process of introducing a national curriculum in mathematics has been a messy one: the original structure has been revised several times, and remains contentious. But throughout the many emergency changes, the philosophy on which the curriculum and its assessment (including the GCSE examination system) have been based has been frequently justified in terms of what was advocated in the Cockcroft report (HMSO 1982) - or more often what that report was deemed to have advocated. Two of the report's oft-repeated demands were: (i) to provide a single curriculum "ladder" up which *all* pupils climb - but at different rates; and (ii) to respect the needs of "the bottom 50%" by designing the initial steps on this ladder, and the associated assessment, "from the bottom up".

The Cockcroft report included no summary list of recommendations on the explicit grounds that "the teaching of mathematics must be approached as a whole" (HMSO 1982, para 809). While one may agree with the reason given, the lack of an "Executive summary" made it too easy for the report to include proposals in different places which were in conflict with each other. Points (i) and (ii) above are therefore this author's attempt to summarise what came to be important principles in the ensuing years, and to capture as fairly as possible the report's remarks concerning the "Foundation list" of topics to be covered by all pupils (Chapter 9), and the structure of a possible single examination system (Chapter 10).

The requirements (i) and (ii) above may help to explain two distinctive features of the resulting English curriculum with regard to "proof" - the first of which is an apparent nervousness about embedding "proof" clearly within ordinary instruction. Pythagoras' theorem, for example, is deliberately listed so as to exclude any reference to its proof: "understand, recall and use Pythagoras' theorem" (NC 2000). Recent revisions have introduced some more encouraging requirements ("they begin to use deduction to manipulate algebraic expressions"; "as they encounter simple algebraic and geometric proofs they begin to understand reasoned arguments" (NC 2000)). But it remains unclear how these "understandings" are expected to grow in the subjective mental soil which is all that is often available - thanks largely to the second distinctive feature of the curriculum discussed in the immediately following paragraphs. Pupils who are not expected to master proofs of standard results, and who are trained to trust their own "convincing arguments" up to age 15, lack the necessary frames of reference with respect to which they might later distinguish between *subjective* and *objective* reasoning - a distinction which is crucial if they are ever to appreciate the essential character of mathematical proof.

The second striking feature is that the curriculum appears to have been based on the mistaken idea that *objective* reasoning in elementary mathematics is the same as *formal* reasoning (based on an axiomatic foundation, together with some attention to logic) - which was felt to be suitable only for a small minority of *older* students. It was therefore decided to re-interpret the idea of "proof" for younger pupils (up to age 14/15 say) by cultivating pupils' own *subjective* "reasoning" and "pattern-spotting" (that is, inference from limited empirical evidence). The wording of the latest official version of the national curriculum is less subjective than it was (NC 2000); but the confused tradition of classroom and assessment practice is still evident in the wording of much supporting guidance. For example:

- Age 8/9: "Pupils show that they understand a general statement by finding particular examples that match it".
- Age 10/11: "They search for a solution by trying out ideas of their own".
- Age 12/13: "They draw simple conclusions of their own and give an explanation of their reasoning".
- Age 14/15: "Pupils are beginning to give mathematical justifications".

In English colloquial speech references to "my reasoning" are almost always inductive rather than deductive, highlighting some presumed *inference*, and bestowing colloquial explanatory power on a personal hunch. The national curriculum and its associated documentation consistently uses words such as "reasoning" and "argument", but with a sleight of hand which changes their meaning at will from *subjective* (for most pupils at all ages) to *objective* (for a minority of pupils, in a very limited number of settings, at a much later age). However, there is now increasing recognition that something is wrong with the presumption that *objective* mathematical proof might somehow evolve naturally, by miraculous transmutation, from its apparent opposite: "It is commonplace in mathematics [education] to present proving in a hierarchy of levels in which the empirical precedes the deductive. This paper questions the assumption that this is a matter of development from the former to the latter " (Hoyles 1998).

The reluctance to embrace a *formal* approach to school mathematics is understandable. But this reluctance was somehow transferred from concerns about *formal* proof to the whole notion of *objective* proof. This was a mistake, since the objective character of elementary mathematics can (and should) be made accessible to all pupils, without ever adopting an inappropriately formal structure.

Mathematicians were themselves partly responsible for this confusion, in that in the 1970s they fell into opposing factions: some sought to re-interpret school mathematics along axiomatic lines (according to some modern parody of Hilbert), while others continued to insist on the primacy of intuition and meaning (in the spirit of Poincare and Thom).

Perhaps the leading recent member of the second group was the late Hans Freudenthal. His inspirational writings in mathematics education devoted surprisingly little space to an explicit analysis of the role of proof. In his monumental 680 page book (Freudenthal 1973) the issue was addressed in the most enigmatic, and the shortest, chapter - just 8 pages long. And in his more detailed (Freudenthal 1983) the subject of proof remained entirely implicit.

Freudenthal and his followers were reacting to what they saw as the misguided formalism of the defenders of the "new math". They therefore focused their efforts on repudiating what they judged to be a false philosophy, and on inventing a convincing alternative to the pseudo-axiomatic approach which dominated the new math era. Freudenthal took it for granted that his readers understood the objective character of mathematical knowledge, and the responsibility of the mathematics teacher to lead students - through engaging with the world of their experience - to an understanding of the higher mathematical realm: "Above all other mental exercises, mathematics has the advantage that with each statement you can decide whether it is right or wrong" (Freudenthal 1973, p. 147). Learning how to "decide" was assumed to be an integral part of the overall programme: "… people have protested to me: 'Eventually the pupils will have to learn the clear and rigorous difference between mathematics and the real world'. I answered: 'You are right if you aim at teaching mathematical rigour, and wrong if you are defending teaching unrelated mathematics' " (Freudenthal 1973, p. 152). What he and his followers rejected was the idea that this transformation could be magically achieved by foisting on students a fake axiomatic approach: "You are … wrong if you advocate teaching ready-made axiomatics" (Freudenthal 1973, p. 152).

The idea that beginners should be spared the objective character of elementary mathematics is not the only error. There is also something seriously wrong with the notion that cultivating beginners' own *subjective* "reasoning" leads on naturally to the notion of *objective proof*. Good mathematics teaching draws on students' own reasoning to encourage discussion and reflection; but it does so in order to challenge and to refine that reasoning, to provide an objective alternative, and to show all students that their own "reasoning" is only useful insofar as it transcends the merely subjective, and learns to respect the criteria for *objective* proof.

This reluctance to incorporate *objective* mathematical proof in some suitable form from the earliest years mistakenly assumes that "the bottom 50%" can master the mathematics they require without ever needing to think abstractly, and that they can thus be spared the luxury of proof. The Cockcroftian assumption that everyone should follow the same initial path then implies that there can be no place for objective proof until the majority of pupils have been left behind - after the age of 14 or so.

The educational instincts of the post-1968 generation were naturally child-centered. Piaget and others emphasised the importance of the child's experience and development of language, and underlined the need always to start from what the child knows. Much of this is now accepted as a truism. But the associated theories concerning "natural stages of development" encouraged an inflated 'romantic' respect for the child's primitive "knowledge"; and this in turn tended to undermine the 'classic' assumption that teachers should lead all students to the promised land of "official mathematics". "Many issues which divide English mathematics teachers into opposing camps seem to arise from our inability simultaneously to keep hold of two complementary ideas in creative tension. For example, we appear unable to grasp the crucial link between the first ('romantic') part, and the second (ultimately more important 'classic') part of the principle that mathematics teaching should start out from where pupils are at, but that this has to be done with the clear objective of exploiting such 'child-centred' *beginnings* in order to achieve important 'content-oriented' *goals*, such as establishing a conceptual platform which is sufficiently strong to ensure that all pupils progress to master proven and important standard methods" (Gardiner 1998, pp. 359-360).

As more children stayed on at school, it became clear that the path from the average young child's intuitive understanding of the concrete world to mastery of the abstract realm of artificial mathematics was far from smooth. A re-evaluation of didactical approaches to key topics, of textbooks and examination structures, and even of the accepted goals for school mathematics was clearly in order. Unfortunately education became politicised, its official ethos changed from "society's cultural duty" to "delivering pupils' rights", and the focus shifted away from what should be taught and onto the pupil as consumer. All of which may help to explain the apparent ease with which some countries came to question whether abstract mathematics was still appropriate for most pupils, and to attribute inflated status to the child's subjective judgement.

Traditional school mathematics sought to challenge pupils' subjective "reasoning", and to supplement it by providing a discipline of standard routines from arithmetic, measures, algebra and geometry. Despite the drawbacks, this repertoire of procedures gave many students an (admittedly passive and authoritarian) understanding of the objective character of mathematics, and dramatically extended the range of their effective action. In contrast, the approach recently adopted in England has had the effect of misleading even the better students into imagining that they are free to adapt the procedures of elementary mathematics as they choose, and that it is more important that they "reason" with confidence than that they reason correctly (see (Gardiner 2003a, 2003b) and the three illustrative examples in section 1.4).

1.2 The dilemma

"Proof" still had to be accommodated somehow, if only because it seemed to be important to mathematicians! However, once the *objective* character of elementary mathematics had been misconstrued as being identical with a *formal* axiomatic treatment, one faced a dilemma - since the idea of a formal axiomatic treatment was out of the question.

Precisely why the dilemma was "resolved" in the way it was remains unclear. It may be that the Cockcroftian principles (i) and (ii) (section 1.1) encouraged the optimistic fusing of the colloquial and the mathematical uses of the word "reasoning", in the hope that more pupils might make more progress if all were encouraged initially to develop "their own (subjective) reasoning". Unfortunately, the mental set of many current 18 year old undergraduates suggests that there are serious dangers in effectively encouraging early "subjective reasoning" in the absence of an objective yardstick. The kind of "reasonings" that are thereby reinforced are often worryingly subjective - even anti-mathematical. By concentrating initially on students' own subjective reasoning in the absence of the discipline of objective mathematical proof (in some form), students have been abandoned to their own unstructured idiosyncratic methods, and have learned to "get by" as best they can with no logical support. When faced with the simplest problems, our best students may manage to structure their own methods; but those outside this tiny elite are condemned to insecurity even in such trivial instances as the *Two cyclists* (section 1.4). And *all* students are likely to come unstuck when the going gets slightly tougher (see *Tom, Dick and Harry* and *Simultaneous equations* in section 1.4).

The results among 18 year old school leavers are scarcely surprising; but they are the precise opposite of what was promised. First year university students, including those majoring in mathematics, increasingly need remedial support; but, in the absence of any common understanding of the objective logical character of mathematics, it is often difficult for those providing that support to explain what is wrong with students' private methods, and hence to help them make progress.

1.3 Towards a possible alternative

The issues discussed in sections 1.1 and 1.2 raise the obvious question of what alternative approach might be more effective. The remainder of section 1 looks more closely at the uncomfortable reality of the current situation in England. But in section 2 we outline a possible alternative, whose aim is to develop the logical capacity of all young children, while remaining faithful - in a simple way - to the spirit of mathematics from the earliest years.

We propose that, from the very beginning of schooling, each major new method or technique should be developed flexibly and orally, with an appropriate measure of practical or exploratory work; but that *the final synthesis* of each method should be routinely *summarised in the form of a short standard written protocol*. These standard protocols should be designed so that any application of the method can, if needed, be presented in a standard form as a sequence of statements, with one statement per line, starting from what is given and ending with what is sought, and such that each line follows naturally and unpedantically from the line before, or from basic facts which are presumed "known".

The approach should be neither axiomatic nor formal. It is rather an attempt to provide students with a sequence of structured formats which make possible the "local organization" of solutions to individual problems relative to some restricted body of accepted facts (Freudenthal 1973, p. 150-151). The underlying pedagogy, and the way each topic is introduced, may remain child-centred; but if ordinary students are to make genuine progress in mathematics, almost all need standard templates of this kind to provide a framework (a) within which their solutions can be presented and checked, (b) by means of which they can be expected to organise a sequence of steps in a way that makes plain to the reader the validity of the final conclusion, and (c) through which their understanding of proof, and their acceptance of responsibility for identifying and correcting errors can mature. The final protocols would serve as a standard format, or common language, for presenting solutions of a particular kind, in much the same way that shared use of correct mathematical terminology sharpens classroom communication. They would also constitute a continuous thread of simple examples of *objective* proof, which slowly but surely convey the essential objective character of the subject. (If videotapes of typical lessons can be trusted, such protocols may well be standard didactical practice in many non-English speaking countries.)

Many of these examples would traditionally be classed as mere "calculations", rather than proofs, but their logical structure is identical to that of "proofs" in the narrower sense. And it is only by admitting such simple examples to the Pantheon of proof that one can aspire to a curriculum, which provides pupils with simple models to serve as objective yardsticks in the struggle to transform their own subjective reasoning into something genuinely mathematical.

The development of such templates needs to recognise the tendency for such frameworks to become an end in themselves, rather than a means to the higher ends of validating correctness and helping students to reflect on the nature of objective proof. This tendency for templates to degenerate led certain wise observers and practitioners to applaud the call (NCTM 1989) for reduced attention to the two-column proof. However, some soon "began hearing people in education claim that proof was an obsolete topic for school geometry … . More than a few teachers were saying, 'We don't do proofs anymore'. Proof had already been eliminated from the low tracks of geometry; it was now about to disappear at every level" (Cuoco 2003, p. 783). Yet if it is true, as we contend, both that students need templates for thought, and that templates have a tendency to degenerate, then the response cannot be to deprive students of these frameworks for thinking. Rather we must choose the standard protocols with care, and recognise the need to develop a professionalism among mathematics teachers which continually reviews and refreshes the didactical basis for whatever templates we may use.

Our proposal accommodates two very different aspects of proof in elementary mathematics: one is profound, the other is mundane - but both are important.

The first aspect is logical and methodological - namely the fact that proof (whatever form it takes) provides mathematics with a procedure which transcends the merely subjective, and which therefore makes it possible to aspire to objective truth. This procedure does not claim to eliminate error; but it combines a formal style and layout designed to make errors transparent, so making it relatively easy for students to take responsibility for identifying and correcting their own errors, in a spirit of openness which invites further public scrutiny.

The second, more mundane, aspect is the particular outward form by means of which we seek to implement this underlying goal: that is, the particular protocol, or form of "book keeping", which we adopt in order to make the whole process transparent, and through which students demonstrate their acceptance of the fact that responsibility for any errors rests with them.

Attempts to achieve a more profound goal are often frustrated by the lack of a suitable mundane frame of reference. Learning one's tables is scarcely a higher-order skill; yet without it, many interesting problems remain out of reach; and mundane hand-eye coordination of pencil, ruler and compasses plays an essential role in mediating the logical exactness of ruler and compass constructions to the mind. In the same spirit, when seeking to devise a long term strategy for teaching proof, it is crucial not only to adopt a broader than usual interpretation of objective proof within the context of elementary mathematics, but also to devise suitable mundane "templates" to provide the unconscious frame for students' thinking, within which the higher goal of objective proof might be explored and mastered. Identifying and simplifying such templates will require an intense programme of design, experimentation, review and refinement.

1.4 Elementary examples

The three examples in this section indicate the extent of the need for change in England, by revealing the actual level of performance of some of our most successful high school graduates - students whose mathematical "success" at school level was achieved the help of restrictive templates to discipline their thinking and their calculation.

It is increasingly recognised that, while one would like all students to think carefully about every problem and to use what they know to respond appropriately, this is only possible for most students if we ensure that they first achieve mastery of the relevant techniques, leaving them free to focus on the particular problem in hand (Barnard 1999, 2002). Similarly, if we want students to present their solutions in a logical form and to check their correctness, they first need some robust standard framework within which to work. Our proposed approach rejects the notion that mathematical proof can be quietly sidelined into some corner of the curriculum intended for a small minority of enthusiasts. Children's earliest experiences of school mathematics - through counting, place value and calculations with positive integers - make clear the *exact* nature of the subject. These experiences of exactness and precision need to be used to give all children a lasting insight into the essential character of mathematics. Applications of elementary mathematics (to measures and practical problems) may introduce the ever-present reality of approximation; but this should not undermine the central message of the objective character of calculation and of exact reasoning within mathematics.

Thus our eventual goal is to indicate how the notion of objective proof might be embedded as a continuous thread within a curriculum for all, by re-interpreting the traditional idea of proof so that it informs and inspires ordinary school mathematics from the very beginning. The initial encounter with each new topic or theme may still emerge from personal experience; but it would then be routinely transformed - through shared analysis - into an objective synthesis which transcends any initial subjectivity. Thus the didactical approach to each standard topic or technique (in counting, measures, calculation in arithmetic or algebra, geometry, etc.) should be such as to build towards this final synthesis. And part of such a synthesis will often be the formulation of a mundane standard format, or template, for presenting all calculations of a particular kind (see section 2).

The function of such a template is two-fold. First, it provides a standard format within which each student's polished solutions are to be presented - not as part of some religious ritual, but in order to lay out clearly the sequence of steps used, so allowing ordinary students to identify and correct their own errors. Second, it provides a framework which can help students to think more clearly about what is needed when they are confronted by harder problems, and so extends their reach and power beyond what might otherwise be possible.

The three examples below are truly shocking. But they are part of a general trend which shows what can go wrong when one places undue reliance on subjective "reasoning" (see also (Gardiner 2003a, 2003b)).

The three problems were given to the 76 first year students (aged 18+) who were present at the final lecture of a first semester university mathematics course. These students had all achieved apparently good grades (A or B) in the school-leaving mathematics examination taken by less than 10% of the cohort, and who were majoring in mathematics at a leading English university. Students were given 22 minutes to solve the three problems. (Most students probably had calculators available in their bags, and their use was not explicitly forbidden. However, the instruction to present clear reasoning, and to cross out errors while leaving them legible, seems to have had the effect that relatively few students used a calculator.)

The first problem illustrates the underlying issue.

The two cyclists. Two cyclists cycle towards each other along a road. At 8am they are 42km apart. They meet at 11am. One cyclist pedals at 7.5km/h. What is the speed of the other cyclist?

This is a problem in elementary arithmetic, which is entirely appropriate for moderately able pupils aged 10-12. Its solution should have taken no more than one minute; but most students appeared to take five minutes or more. The task of extracting and combining simple information efficiently and reliably proved surprisingly challenging. Of the 76 students present, 26 (34%) failed to reach the answer 6.5km/h, and several more made serious errors before a second - ultimately successful - attempt. (Five of those deemed to have failed realised the need to calculate $19.5 \div 3$; however, since three of these students evaluated this incorrectly, the two attempts which left the answer in this form were deemed incomplete.)

One possible inference is that what we often excuse as "mistakes" may not be mistakes at all, but are rather a routine and predictable result of what happens when students reach their (very low) threshold for "information overload". Many of these students would appear to have rarely been required to take full responsibility for tackling the simplest multi-step problems. If this is indeed the case, then one can be fairly sure that such errors will occur whenever these students have to coordinate two or more facts or methods - no matter how simple these facts or methods may be.

The successful solvers *all* utilised the obvious method. But they had to devise this for themselves, inventing their own layout and style of presentation! While 50 students were successful, 26 were not: the lack of any standard format for presenting solutions to such problems can only have increased the level of demand - both for those who failed and for many of those who succeeded - in what should have been a completely trivial problem.

While all three problems illustrate the worrying level of achievement currently attained, their main purpose here is to indicate the need for a more structured approach to calculation - designed to help students present steps within an agreed standard framework - as part of a long term strategy for teaching the nature of local deduction.

The second example deals with "rates", and is of a kind which is important in mathematics and in science. Such problems used to be standard, but have been completely neglected in recent years. Nevertheless the problem is still entirely appropriate for able students aged 15-17, who one would expect to respond intelligently, even if they ultimately fail to obtain a complete solution. The third example is an unfamiliar, but perfectly accessible, pair of simultaneous equations, which students majoring in mathematics should again be able to handle sensibly, even if a complete solution eludes them.

Tom, Dick and Harry. Tom and Dick take 2 hours to complete a job. Dick and Harry take 3 hours to finish the same job. Harry and Tom take 4 hours to finish the job. How long would all three take working together?

Simultaneous equations. Solve the simultaneous equations:

$$x^2 - y^2 = -5$$
$$2x^2 + xy - y^2 = 5.$$

Not one of the 76 students solved the second problem and just three students solved the third problem. However, the scripts themselves are more instructive than this crude summary of successes and failures. What is most revealing is the remarkable uniformity of the crass error in *Tom, Dick and Harry* (which emphasises the need for standard templates as an aid to student thinking), and the bewildering variety of almost random approaches which students adopted "on the hoof" in *Simultaneous equations*, with each student creating a unique concoction of errors, and a uniquely personal wrong answer.

In *Tom, Dick and Harry*, ten students wrote nothing more than some abbreviated version of the given data; four students produced moderately intelligent estimates (such as 4/3 hours), though without admitting that they were guessing. Of the remaining 62 students, 58 wrote the equations: T+D=2, D+H=3, H+T=4, and proceeded to try to solve them, while four students wrote some unexplained variation on this theme (one of which looked interesting, but remained incomplete). Not one of these students ever fully realised the meaninglessness of what they had written: some efforts petered out; others made simple arithmetical or algebraic mistakes. 30 students ground their way to the conclusion "combined time = 4.5 hours"; but only eight of these remarked: "This must be wrong"; and not one felt any apparent obligation to identify their error, or to do anything about it.

In the third problem just three students derived the two answers more-or-less correctly (e.g. omitting only to justify division by $x+y$ or by $x-y$ at some point). One student simply wrote down the two solutions without any explanation - presumably spotted by trial and error. Three students obtained the "answer" $x=\pm2$, $y=\pm3$ without realising the need to pair each value of x with a single value of y (the two possible values of x having been simply substituted into the first of the given equations to obtain $y^2=9$). Five students obtained $x=2$, $y=3$ only (some by ignoring the negative root of $x^2=4$, but some via other algebraic errors and guesses). One student obtained the single answer x=-2, y=-3 (by making the remarkable step $y^2 + xy - 15 = (y+3)(y-5)$ to infer the solution y=-3). Three other students got as far as showing that $2y = 3x$, with the usual algebraic oversight of dividing by $x \pm y$ without justification. The remaining 60 students made a bewildering variety of algebraic errors of varying degrees of crassness.

The student scripts for these three problems reinforce the impression gained from (Gardiner 2003a, 2003b): our fear of teaching and practising mathematics within a thoughtful framework of interconnected rules and standard templates would seem to guarantee that only a tiny handful of students emerge at age 18 with any understanding or mastery of elementary mathematics. Though there are clear pedagogical dangers in using such rules and templates blindly and inflexibly, the danger of abandoning (even fairly able) students to operate without such a supporting framework seems to be far greater.

This interim conclusion has one consequence which is directly relevant to the issue raised in the next section. If, on the one hand effective mass education in mathematics requires a supporting framework of rules and standard templates, while on the other hand rules and templates have a tendency to degenerate, then school mathematics should concentrate on mathematical topics (a) which are recognised as being of central importance, and (b) which are sufficiently rich in connections and problems to make it possible to counteract any tendency to degeneration. The second of these conditions is especially important whenever the material is to be assessed through formal written examinations. (It may be that, at school level, certain traditional domains - such as arithmetic, measures, percentages, fractions and ratio, euclidean geometry, algebra, analytic geometry, trigonometry and calculus - can be taught and assessed in ways that satisfy these requirements, while more recent topics - such as sets, transformation geometry and data-handling - cannot.)

1.5 The case of discrete mathematics

For the last 10 years or so, discrete mathematics has been available in England as an alternative to mechanics and statistics for the 10% of students who continue studying mathematics at age 16-18. The resulting experience suggests that discrete mathematics may be an example of a domain which, at the high school level, fails the second requirement (b) in the previous paragraph.

The most basic theme within discrete mathematics is that of *counting*. But though the associated sum and product rules for counting can be understood in the simplest cases, their use in harder examples requires a flexibility of thinking which makes intellectual demands that stump many students. Problems involving "permutations and combinations" soon separate a class into the few who can, and the many who cannot, see how to begin. And it is not easy to move beyond the common sense version of the inclusion-exclusion rule (filling in numbers on a Venn diagram) to more mathematical applications. This may explain why the natural theme "counting" does not really feature in discrete mathematics syllabuses (binomial coefficients are defined as part of algebra at this level, but they are not used for counting).

In the 1980s numerous reports advocated discrete mathematics as being more appropriate than calculus for many college students. Since then the claims for discrete mathematics have been extended to high school level (NCTM 1989, 1991), though without indicating which topics "of lesser importance" should be discarded.

We focus here on two claims made on behalf of discrete mathematics, which are relevant to our consideration of "proof": "Discrete mathematics fosters *critical thinking* and *mathematical reasoning*" (NCTM 1991, p. vii); and "In grades 9-12, the mathematics curriculum should include topics from discrete mathematics so that all students can … develop and analyze algorithms" (NCTM 1989, p. 176). In reality, there seems to be a marked temptation to skirt round matters of proof and analysis in discrete mathematics - even at undergraduate level. At age 16-18 the temptation is all but irresistible. Thus the English experience of the last 10 years suggests the precise opposite of the two bold claims above.

The example of "counting" illustrates two reasons why discrete mathematics is inappropriate for most beginners. Problems in discrete mathematics tend to be either mindlessly routine or impossible! Non-trivial looking problems are mostly too hard; so they can be included only by restricting attention to a tiny number of artificial stereotypes, for which students are taught to apply rules blindly, without any real expectation of understanding.

For example, the first serious chapter in (Hebborn 2000) - the official textbook for the most popular discrete mathematics syllabus - introduces "modelling with graphs". The second serious chapter then introduces a sequence of "algorithms on graphs", beginning with *Kruskal's algorithm* for a minimum spanning tree:

"**Step 1** Sort the edges in ascending order of weight.
Step 2 Select the edge of least weight.
Step 3 Select from edges not previously selected the edge of least weight that does not form a cycle together with the edges already included.
Step 4 Repeat step 3 until selected edges form a spanning tree" (p. 52).

No attempt is made here - or anywhere else in the text - to encourage the student (or the teacher) to consider the question as to why this should always produce a *minimum* spanning tree. Hence the impression conveyed is that a "greedy" algorithm is automatically globally optimal.

The very next section presents *Prim's algorithm* for a minimum spanning tree. Again no proof is offered. The algorithm is interestingly different from the *Kruskal algorithm*, and may produce a different output. Yet the comparison of the two algorithms is restricted to a single comment, concerning the need in *Kruskal's algorithm* to sort the edges in ascending order by weight, and the need to check whether a new edge creates a circuit. However, since all the examples students ever meet are small, they are unlikely to appreciate the significance of this remark.

This overt neglect of "critical thinking" and of "mathematical reasoning" permeates the whole text - and presumably the way the material for this popular syllabus is taught and assessed. The first edition of the text (Hebborn 1997) had included outline proofs; but perhaps because there was no easy way to test the proofs in the final exam, students and teachers tended to ignore them, and the text has since been streamlined. This "functional" approach ignores the fact that proofs in mathematics - and especially in discrete mathematics - are often essential for alerting the learner (and the teacher and textbook author!) to hidden subtleties. For example, both editions fail to note that Kruskal and Prim apply only to *connected* graphs, so the first step should be to check whether the graph is connected (if the graph is small and is represented graphically - rather than being specified by a matrix - this move may seem redundant; but under such conditions, the algorithms themselves are unnecessary). Such oversights are typical - reducing the mathematics to a set of half-comprehended rules. For example, in (Hebborn 2000) the given "algorithm" to test for the non-planarity of a graph (p. 72) begins blithely "Step 1: Identify a Hamiltonian cycle in the graph"! In discussing Eulerian graphs, an outline indicating the necessity of the standard condition is presented, and the impression is then given that the condition has been proved to be sufficient (p. 89). And when discussing the algorithm for finding a maximum matching, the words *maximal* and *maximum* are used interchangeably (e.g. p. 207-208), ignoring the crucial distinction - for beginners as well as for mathematics - between a *maximum* matching (with the maximum possible number of edges) and a *maximal* matching (one which cannot be extended).

If a mathematical technique is to be understood, and used intelligently, then it should ideally be mastered together with some suitable version of its proof. The main problem with discrete mathematics is that, if students are to have any chance of understanding such a proof, they must feel completely at home with the underlying universe of mathematical objects. And, though the universe of "all finite graphs" is elementary in the sense that it is "discrete", it is considerably more elusive than the universe of familiar numbers.

Thus the student is reduced to implementing uncomprehended algorithms, in unrelated and stereotyped problem situations, like some very slow and unreliable computer. Discrete mathematics could easily give rise to a supporting framework of "standard templates"; but at this level it lacks the necessary richness of problem material, and the logical connections between topics are perhaps too subtle. Hence it neither manages to cultivate flexibility and mathematical reasoning, nor to avoid degeneration.

As a result the assessment, and hence the teaching, of this material has degenerated. Many students (and their teachers) "like" the discrete mathematics options, because the material is felt to be "different", and because the assessment items are so predictable that an industrious student can score high marks. But in the hands of most teachers there is little chance of any "critical thinking".

Such degeneration is in no way restricted to discrete mathematics. There is a widespread reluctance in mathematics education to recognise, when revising curricula, that the impact of assessment procedures on day-to-day classroom practice is a key factor in the ultimate success or failure of reform. It is disingenuous (and, given increasing public interest in "accountability", probably futile) to argue for the abolition of formal written assessment, or for it to be incorporated within instruction. So we need an approach to core content and its assessment which encourages quality teaching and learning of mathematics.

One of the advantages of developing suitable standard templates for key topics from sufficiently rich problem domains within school mathematics is that they provide an obvious way of ensuring that "proof" is routinely assessed - and hence routinely taught.

2. From calculation to proof: recognising calculation as "proto-proof"

We take a (proto-)*proof* to consist of:

- any sequence of statements, each of which is clearly formulated and clearly laid out, and is either self-evident from standard known facts or from the structure of the argument presented, or is clearly justified in terms of previous steps, or known results;
- with the first statement being known to be true (or being a clearly identified hypothesis which will be disproved), and the last statement being that which was wanted.

This broader-than-usual conception of proof is taken from Gardiner (1999b), and is echoed by Barnard (2002): "We shall take the term 'mathematical proof' to mean a hierarchy of links between givens and a concluding statement, where a 'given' is something that is assumed (explicitly or implicitly) without relating it to anything more primitive. A basic ingredient in the building of such links is the manipulation of mathematical statements" (p. 121). The formulation is aimed at mathematics educators, and is not intended (at least not initially) for students. It applies to, but extends far beyond, those areas traditionally associated with proof (such as euclidean geometry); for it has been worded so as to apply to any mathematical argument or calculation involving at least two steps. As preliminary examples we offer:

Example 1: Evaluate 13 + 26 + 37 + 44 as efficiently as possible.

$$\begin{aligned} 13 + 26 + 37 + 44 &= (13 + 37) + (26 + 44) \\ &= \quad 50 \quad + \quad 70 \\ &= 120. \end{aligned}$$

Example 2: I buy 7 apples and get 16p change from £1. What does each apple cost?
Suppose each apple costs x pence.

$$\therefore\ 7x + 16 = 100$$
$$\therefore\ 7x = 84$$
$$\therefore\ x = 12.$$

Example 3: Multiply out $(a\text{-}b+c)(a+b\text{-}c)$ as efficiently as possible.

$$(a-b+c)(a+b-c) = [a - (b-c)][a + (b-c)]$$
$$\therefore\ (a-b+c)(a+b-c) = a^2 - (b-c)^2$$
$$\therefore\ (a-b+c)(a+b-c) = a^2 - b^2 - c^2 + 2bc.$$

In the context of "learning to prove", these examples need to be embedded in a classroom setting where the standard template (or some alternative) which underpins each example has already been made available as a natural frame of reference. In particular, each standard format needs to be developed, practised and internalised by writing out solutions to lots of simple problems, before it can be used to extend the range of problems which can be solved successfully by all students.

Such problems could of course be tackled and solved by individual pupils using "their own reasonings". One may even hope that most such approaches would arrive at the "right answer". But some would inevitably be flawed in some way, and many would lack clarity. Requiring students from time to time to present their solutions in the agreed line-by-line format of a standard protocol would help to make the inner logical structure explicit.

The proposed approach is scarcely sophisticated. Yet honesty compels one to concede how much work would be needed to implement such an approach on a wide scale. Proof is a way of organising calculations within a given framework - whether with numbers, with symbols, with geometrical entities or with logical propositions. There is no escape from the fact that this presupposes two things. First, a social discipline which allows the teacher to insist on a measure of conformity in adopting and using mundane frames of reference and deductive principles, which are common rather than idiosyncratic, and which are perceived not as shackles, but rather as the soil within which creativity can flourish. Second, a three-fold appreciation on the part of the student that mathematics is *exact*; that if one looks at things in the right way, one can expect answers to be comprehensible - and frequently simpler than expected; and that proof, or exact calculation, offers the only reliable way of harvesting this simplicity.

To echo what we wrote earlier, the problem - at least with English 18 year olds entering university to study numerate subjects - would seem to be not that students have some incidental difficulty in adhering to and implementing such common procedures, but rather that they have no clear conception of the deductive character of calculation, and so do not see the need for working within a standard framework which might allow them to take responsibility for and to evaluate the correctness of their own solutions. However, if we are to teach mathematics at school level, such difficulties need to be understood and faced.

Example 1 and Example 3 illustrate the pedagogical advantages of using "contrived calculations" to counteract the incomprehension referred to in the previous paragraph. Problems involving "real data" often encourage students to hack through every calculation from the beginning, without ever internalising the routine expectation that what may at first appear complex is generally simpler than it looks, and can often be analysed and comprehended by the human mind. If one wishes to encourage structured thinking, with solutions laid out in a standard way to make the internal logic clear, then the numbers need to be chosen precisely to reward and to cultivate the kind of *irrational optimism* without which the beginner sees no reason to look beneath the surface to identify the hidden structure in a problem. In the absence of this instinct for sense-making, students resort too easily to unstructured, and hence error-prone, calculation, or to apparently random moves. Effective mathematics education actively cultivates such "irrational optimism" in students, so that they learn to look for - and expect to find - helpful structure just below the surface.

None of the examples is what is normally understood by a "proof". Yet each provides pupils with a clear yardstick which can help them refine "their own (subjective) reasonings" into mathematical proof. In the first example - as with most calculations at this level - the goal is to reduce the calculation to a short sequence of indisputable steps, which removes all doubt, while the deductive character of each step remains implicit. The second example adopts a standard approach and layout which makes the underlying logical structure explicit: each line represents a new step, and the connections between successive steps are established via the use of the "therefore" symbol. The third example is an algebraic variation on the first - avoiding the error-prone strategy of multiplying out all nine terms before cancelling and collecting, seeking instead to reduce all calculation to the two well-known identities for $(a+b)(a-b)$ and for $(a+b)^2$.

There is another important aspect of Example 1 (at age 6/7), of Example 2 (at age 12/13) and of Example 3 (at age 15/16). Pupils' own calculations at each level are often inefficient, even when successful. If they are ever to appreciate the decisive, objective character of the underlying steps, it is important to have a standard format which allows one to summarise those calculations which can be presented simply in short objective *written* form - so that the advantages of re-grouping in Example 1, of the standard approach to Example 2, and of recognising the difference of two squares in Example 3 can be clearly grasped, and the indisputability of the answer recognised.

These examples should be seen as simple instances within an extended sequence, which systematically exploits children's early appreciation of "objective" reasoning (reinforced, as Piaget showed, by experience of the world and by the use of language) to help them develop over time a clear idea of what is meant by deductive proof, and its marked difference from subjective reasoning.

Early examples from the realm of calculation - whether with numbers or with symbols - are sufficiently simple that the sequence of steps can usually be chosen so that each line follows naturally from the previous line, with no need to appeal to interim conclusions or external results. The justification for each step is then clear *from the ordering of the steps*. Thus, while each step should be explained verbally when presenting such a proof, there is no need to require that it be written out explicitly. Moreover, with arithmetical calculations, or with linear problems, each step is reversible; thus, while there may be good psychological reasons to insist that the answer be checked, it would be pedantic to see this as part of the proof structure at this level.

However, the advent of problems involving squares or square roots leads to steps which are definitely not reversible. There is then no escaping from the need to confront - in some form - the fact that deduction yields a list of *candidate* answers, rather than guaranteed answers. At this point - if not before - it becomes clear that each step in a proof sequence may need to appeal to more than just the immediately preceding step, and that where this is needed, the justification (for example, when eliminating certain candidate values) has to be made *explicit*.

More sophisticated proofs routinely involve steps which can only be justified by explicit reference to clearly identified *external results*. This is especially true of euclidean geometry, where in each given problem one looks for ways of exploiting one of a relatively small number of standard external results (the angle sum of a triangle; criteria - vertically opposite, alternating, etc. - for two angles to be equal; isosceles triangles; the SAS and SSS congruence criteria; Pythagoras' theorem; formulae for the area of a triangle; the sine and cosine rule; angles in the same segment; etc.). Euclidean geometry may provide the richest accessible example; but the need to identify and apply some standard external result is typical of mathematics and characterises the solution of many beautiful elementary problems (see Gardiner 1997, 1999a).

Implementing such an approach in a manner that avoids degeneration will not be easy. But the present situation is unacceptable. And moves to re-interpret mathematics-for-all in terns of mere "numeracy" may yet make things worse! So it is essential for committed educators and mathematicians to work together to devise, implement and refine strategies which reflect both the discipline of mathematics and the way ordinary students learn.

References

Barnard, T. (1999). A pocket map of algebraic manipulation - Leicester, UK: The Mathematical Association.

Barnard, T. (2002) Reducing mathematics to human size. Research in mathematics education Vol. 4, p. 115-126.

Cuoco, A. (2003). Teaching mathematics in the United States. Notices of the American Mathematical Society Vol. 50, p. 777-787.

Freudenthal, H. (1973). Mathematics as an educational task - Dordrecht, Holland: Reidel.

Freudenthal, H. (1983). Didactical phenomenology of mathematical structures - Dordrecht, Holland: Reidel.

Gardiner, A. (1997). More mathematical challenges. - Cambridge, UK: Cambridge University Press.

Gardiner, A. (1999a). Mathematical puzzling. - Mineola, New York: Dover.

Gardiner, A. (2003a). What does English school mathematics achieve at present? - Birmingham: Birmingham University, School of Mathematics

Gardiner, A. (2003b). What does English university mathematics achieve at present? - Birmingham: Birmingham University, School of Mathematics

Gardiner, T. (1995). Wrong way. Go back! - In: Mathematical Gazette Vol. 79, p. 335-346.

Gardiner, T. (1998). The art of knowing. - In: Mathematical Gazette Vol. 82, p. 354-372.

Gardiner, T. & Moreira, C. (1999b). Proof matters. - In: Mathematics teaching Vol. 169, p. 17-21.

Hanna, G. & Jahnke, H.N. (1996). Proof and proving. - In: A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, C. Laborde (Eds.), International handbook of mathematical education. Dordrecht, Holland: Kluwer, p. 877-908.

Hebborn, J. (1997). Decision mathematics 1. - Oxford, Heinemann.

Hebborn, J. (2000). Decision mathematics 1 (revised edition). - Oxford, Heinemann.

HMSO (1982). Mathematics counts. - London: Her Majesty's Staionery Office.

Hoyles, C. (1998). A culture of proving in school mathematics? - In: J.D. Tinsley & D.C. Johnson (Eds.), Information and communications technology in school mathematics. London, Chapman & Hall.

LMS (1995). Tackling the mathematics problem. London: London Mathematical Society.

NC (2000). http://www.nc.uk.net/index.html

NCTM (1989). Curriculum and evaluation standards for school mathematics. - Reston, VA: National Council of Teachers of Mathematics.

NCTM (1991). Discrete mathematics across the curriculum, K-12. Reston, VA: National Council of Teachers of Mathematics.

Author

Gardiner, Tony, Dr., School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK
Email: a.d.gardiner@bham.ac.uk