Modelling Mathematical Argumentation: The Importance of Qualification†

Matthew Inglis
Learning Sciences Research Institute, University of Nottingham.

Juan Pablo Mejia-Ramos
Institute of Education, University of Warwick.

Adrian Simpson
School of Education, University of Durham.

Abstract. In recent years several mathematics education researchers have attempted to analyse students' arguments using a restricted form of Toulmin's (1958) argumentation scheme. In this paper we report data from task-based interviews conducted with highly talented postgraduate mathematics students, and argue that a superior categorisation of genuine mathematical argumentation is provided by the use of Toulmin's full scheme. In particular, we suggest that modal qualifiers play an important and previously unrecognised role in mathematical argumentation, and that one of the goals of instruction should be to develop students' abilities to appropriately match up warrant-types with modal qualifiers.

The analysis of mathematical arguments produced by students and mathematicians has been a recurring theme in the mathematics education literature. Generally these types of analysis are of two kinds: those that concentrate on the argument's *content* and those that concentrate on the argument's *structure*. This paper fits into the second category. It explores a scheme for describing the structure of arguments, demonstrates the value of examining a core part of that scheme (the modal qualifier) and categorises some of the different forms of argumentation used by mathematicians in realistic proving contexts. These forms are inextricably linked to the extent to which an argument may be qualified or potentially overturned; in highlighting this link the paper emphasises the importance of *reducing uncertainty* in general mathematical argument, alongside the goal of *removing uncertainty* which is considered one of the hallmarks of proof.

1. Toulmin's *The Uses of Arguments*.

Toulmin (1958) advocated an approach to analysing arguments that dramatically departed from existing approaches to formal logic. He was less concerned with the logical validity of an argument, and more worried about

---

the semantic content and structure in which it fits. Indeed, this led to one contemporary describing *The Uses of Arguments* as "Toulmin's anti-logic book" (a description reported by Toulmin, 2001). This manner of analysing argumentation has become known as 'informal logic' in order to emphasise its differences from formal logic.

Toulmin's (1958) scheme has six basic types of statement, each of which plays a different role in an argument. The *conclusion* (C) is the statement of which the arguer wishes to convince their audience. The *data* (D) is the foundations on which the argument is based, the relevant evidence for the claim. The *warrant* (W) justifies the connection between data and conclusion by, for example, appealing to a rule, a definition or by making an analogy.[1] The warrant is supported by the *backing* (B) which presents further evidence. The *modal qualifier* (Q) qualifies the conclusion by expressing degrees of confidence; and the *rebuttal* (R) potentially refutes the conclusion by stating the conditions under which it would not hold. It should be noted that, in any given argument, not all of these roles will necessarily be explicitly verbalised.

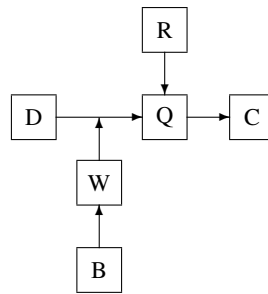These six components of an argument are linked together in the structure shown in Figure 1.

Figure 1. Toulmin's model of a general argument. The argument would read "D, and since W (given B) we can Q conclude C, unless R".

An example of an argument expressed in this form is given in Figure 2. Here, the arguer is suggesting that, during a football match, Hislop was at fault for the goal (C) – a close range header scored whilst he was the goal-keeper (D) – because he ought to have caught the cross from which the goal was scored (W). This is because goalkeepers are expected to be able to catch crosses that are close to them (B). The arguer believes his argument to be likely to be sound (Q), but accepts that the argument could be rebutted if Hislop was fouled as the ball was crossed (R).

In his original work Toulmin (1958) suggested that formal mathematics was one of the few domains of explanation where formal logic – the system he was reacting against – adequately described argumentation structures. However, in a later work, Toulmin et al. (1984) gave an example of a for-
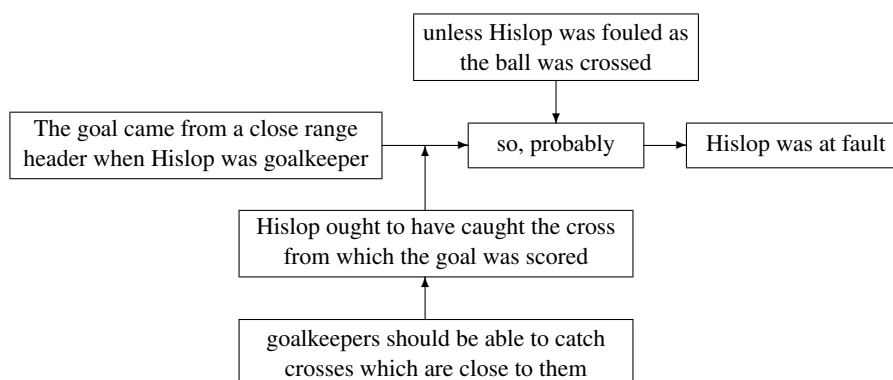
Figure 2. An argument expressed using Toulmin's structure.

mal mathematical argument modelled using his scheme. Further examples of applying informal logic to mathematics have been developed in recent years.

In the field of mathematics education, Krummheuer (1995) started the trend of using Toulmin's scheme by analysing classroom-based mathematical arguments. However, he applied a reduced version of the original scheme, omitting the use of the rebuttal and the modal qualifier, apparently seeing them as irrelevant to mathematical arguments. Although this restricted version of the layout may have been sufficient to provide the level of analysis of the classroom episodes Krummheuer wanted, it is unclear how this omission can be justified in a conceptual framework aimed at the reconstruction of argumentation which may lack logically necessary conclusions.

Most subsequent mathematics education research appears to have followed Krummheuer in using the reduced scheme. This approach has been adopted by, amongst others, researchers studying basic number skills (Evens and Houssart, 2004), logical deduction (Hoyles and Küchemann, 2002; Weber and Alcock, 2005), geometry (Knipping, 2003; Pedemonte, 2005; Pedemonte, in press), and general proof (Yackel, 2001). Indeed, this position appears to have become so entrenched that, in her recent review of research on proof in mathematics education, Mariotti (2006) referred to Toulmin's scheme as a "ternary model".

Authors from other disciplines have adopted a different approach to adapting Toulmin's brand of informal logic to formal mathematics. Aberdein (2005, 2006) and Alcolea Banegas (1998), when analysing formal mathematical proofs using Toulmin's scheme, retained all of its constituent parts including modal qualifiers and rebuttals. In all cases, however, these were filled with psychologically inconsequential values, including qualifiers such as "classically" and "constructively". While these qualifiers may be useful when modelling formal proofs from a philosophical standpoint, they are less rel-

evant for our purposes, as they do not refer to the strength of a person's belief in a conclusion, given the argument's data, warrant and backing.

In this paper, we will demonstrate that, in order to analyse the full range of mathematical argumentation (which we interpret to include both informal reasoning and formal proof[2]), it is important to use Toulmin's full argumentation scheme. Whilst it is well documented that professional mathematicians rely on non-formal arguments as they develop their ideas (e.g. Burton, 2004; Hadamard, 1945; Poincaré, 1905; Thurston, 1994), this paper seeks to explore the kinds of non-formal arguments they use, how they can be analysed, and the potential differences between expert and novice mathematicians in their use of non-formal arguments. We give examples of arguments developed by talented mathematicians which require an analysis involving the elements most often ignored in analysing mathematical argument – the modal qualifier and rebuttal – and we show that these elements are closely linked to a categorisation of the type of warrant that mathematicians use.

2. Method and Participants.

The data reported in this study come from a series of task-based interviews conducted with highly successful mathematics graduates studying for postgraduate degrees. The task, designed as part of a larger study to investigate how successful mathematicians evaluate conditional statements, was based on an adaptation of the materials used by Markowitz and Tweney (1981).

Participants were interviewed alone in a seminar room or private office. The interview began with the participant being given some information, on a A5 card, as shown in Figure 3. When participants indicated that they were ready to proceed they were given the first of the conjectures, and as each conjecture was completed to their satisfaction, they were given the next – each presented on a new A5 card. The full list of conjectures is shown in Figure 3. A mathematical discussion of each of the conjectures, including brief solutions, can be found in the Appendix.

The procedure took the form of semi-structured clinical interviews (in the sense of Ginsburg, 1981), with the interviewer prompting participants when their method was unclear, and asking for clarification as appropriate. Discretion was used by the interviewer in determining whether or not to miss out certain conjectures in order to maximise the use of time. Participants had unlimited amounts of paper available on which to work. As they moved on to subsequent conjectures, they retained access to their previous cards and work. If the participant began to check some examples, or asked whether any examples were available, they were given a further card, shown in Figure 3.

At first sight, one might consider the conjectures to be simple number theory problems which the talented mathematicians would find straightfor-

All the numbers below should be assumed to be positive integers.

Definition. An *abundant number* is an integer n whose divisors add up to more than $2n$.

Definition. A *perfect number* is an integer n whose divisors add up to exactly $2n$.

Definition. A *deficient number* is an integer n whose divisors add up to less than $2n$.

Example. 12 is an abundant number, because $1 + 2 + 3 + 4 + 6 + 12 = 28$ and $28 > 2 \times 12$. However, 14 is a deficient number, because $1 + 2 + 7 + 14 = 24$, and $24 < 2 \times 14$.

Your task is to consider the following conjectures and determine, with proofs, whether they are true or false.

Conjecture 1. A number is abundant if and only if it is a multiple of 6.

Conjecture 2. If n is perfect, then kn is abundant for any $k \in \mathbb{N}$.

Conjecture 3. If p_1 and p_2 are primes, then $p_1 p_2$ is abundant.

Conjecture 4. If n is deficient, then every divisor of n is deficient.

Conjecture 5. If n and m are abundant, then $n + m$ is abundant.

Conjecture 6. If n and m are abundant, then nm is abundant.

Conjecture 7. If n is abundant, then n is not of the form p^m for some natural m and prime p.

Examples.

The first few abundant numbers are: $12, 18, 20, 24, 30, 36, 40, \ldots$

The first few perfect numbers are: $6, 28, 496, 8128, \ldots$

Figure 3. The instructions, conjectures and examples given to participants.

ward. However, while the problems might be accessible, the evidence from the interviews was that the participants found them quite hard - in an intensive hour-long working session, only one participant (David) finished all the problems to their satisfaction.

While the problems were substantial ones for the participants, it is worth noting that the task set for them did not mimic genuine research (nor was it intended to) and, in being in a given context (number theory), was likely to restrict the forms of argument used (e.g. we would expect few visual arguments). Nevertheless, the nature of the problems was sufficient to allow us to begin to analyse the types of argument used by sophisticated mathematicians when faced with genuine problems and, in particular, to examine the link between different parts of their arguments.

Data from six of the interviews from the larger study are reported in this paper. All of the participants were studying in highly ranked UK uni-

versities; five were studying for doctorates and one, Fred, for a masters degree: all had been highly successful in their undergraduate mathematics degrees. Interviews were recorded with an electronic audio-dictaphone, and later transcribed for analysis. The analyses took the form of a multiple case study approach, following the quasi-judicial procedure developed by Bromley (1986).

3. The need for modal qualifiers.

As discussed above, previous work that has applied Toulmin's (1958) argumentation scheme to mathematical reasoning has tended to ignore the role of the modal qualifier and rebuttal parts of arguments. This section demonstrates that appreciating the role of the modal qualifier is crucial to fully understand how mathematicians argue.

When modelling arguments using Toulmin's (1958) scheme it is often the case that certain parts of the argument (most commonly backings and rebuttals) are not explicitly verbalised by the arguer. In line with earlier researchers who have used the scheme, we dealt with this issue by inferring the backings and rebuttals of participants' arguments where they were not explicitly verbalised. Consequently the diagrams reported in the remainder of this paper represent plausible models which account for participants' behaviour and utterances; they are not direct one-to-one mappings from utterance to argument. However, it should be noted that, in the examples given here, very few sections of the diagrams are not directly related to the participant's actual spoken words. The methodological issues involved in using Toulmin's scheme for modelling empirical data are discussed in depth by, for example, Bromley (1986) and Simosi (2003).

A clear example of the use of qualification came from Chris's work on Conjecture 3. After Chris had correctly identified that the conjecture as written was false, the interviewer asked him whether he thought the conjecture would be true if it was modified to read "if p_1, p_2 are prime, then $p_1 p_2$ is *not* abundant". Chris tried two examples (2 & 3, and 5 & 97) to investigate the situation and then remarked:

CHRIS: Since the smallest numbers I could find to put in this equation showed it was perfect and in the larger limit it showed, $p_1 p_2$ was deficient. So it's possible it holds for all p_1, p_2.

INTERVIEWER: Do you think it does?

CHRIS: I think it probably does. But I'm not sure why [LAUGHS]. Yeah, the fact that this is, in some ways, sort of monotonic. In other words, I know that this statement is true for large p_1, p_2; I know it's true for small p_1, p_2; so I feel therefore that it should be true for p_1, p_2 in the middle. Umm, but I might have to do some work to show that.

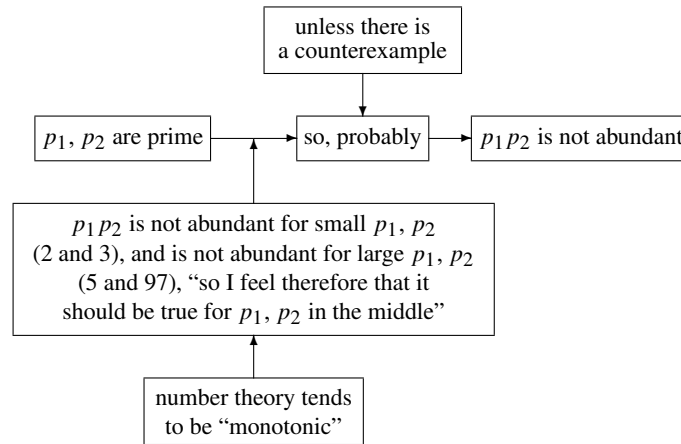The structure of this argument is shown graphically in Figure 4.

Figure 4. Part of Chris's response to Conjecture 3.

Chris says that he thinks it is "probable" that the statement is true, on the basis of two examples and an argument related to monotonicity. He accepts that he has not shown the result formally, but informally he has persuaded himself that the statement is probably true, and does not feel obliged to carry on and produce a formal proof. In terms of Toulmin's (1958) scheme, his argument revolves around a modal qualifier that *does not carry certainty*, instead his uncertainty has merely been reduced.

In this example Chris appears to be fairly confident that the conclusion can be drawn, but there are examples of less certain modal qualifiers. Here, for example, David is asked the same question, about the modified version of Conjecture 3:

INTERVIEWER: If I changed it [CONJECTURE 3] then, to be 'not abundant', what would you say?

DAVID: That would seem more reasonable. Because primes look very deficient.

Compare the different modal qualifiers used by Chris and David. David appears less convinced by his argument than Chris did, felt the need to convert this informal piece of reasoning into a formal proof, and went on to do so. Chris, on the other hand, was sufficiently convinced of his argument that he didn't feel the need to continue, despite accepting that "some work" would be needed to prove the result.

The key point here is that Chris and David used different types of warrants in their respective arguments. These different warrants were accompanied by different modal qualifiers, which fixed their degree of belief in the conclusion of the argument. The approach taken by earlier researchers to use a

restricted scheme without modal qualifiers would fail to highlight the difference between Chris and David and, arguably, could dismiss the forms of argument they use as non-mathematical, despite appearing to be crucial in their development of solutions to the problems posed. We argue that, by using the full Toulmin scheme, we gain access to a wider range of distinctions in mathematical argument.

In the rest of the paper, we highlight a rough categorisation of mathematical argumentation on the basis of *warrant-types*. Whereas a *warrant* is a part of a particular argument, a *warrant-type* is a category of warrants with similar properties. The warrant-types categories have a broadly similar range to the 'proof-schemes' outlined by Harel and Sowder (1998) and consequently this paper adopts some of Harel and Sowder's nomenclature in our description of warrant-types.

However, it is important to note that we are not categorising 'proof schemes', 'proofs' or even 'attempted proofs'. The talented mathematicians in this study are fully aware that some of their arguments do not constitute formal proofs – they are, however, mathematical arguments and are often vital to their solution of the problem. Crucial to the difference between warrant-type and proof scheme is that Harel and Sowder (1998) define a person's proof scheme as that which allows them to "remove her or his own doubts about the truth of an assertion". A proof scheme, then, is about *removing* uncertainty. Warrants from certain warrant-types, in contrast, may only *reduce* uncertainty. This distinction is developed in greater depth in later sections.

4. The inductive warrant-type.

Harel and Sowder (1998) define inductive proof schemes as "when students ascertain for themselves and persuade others about the truth of a conjecture by *quantitatively evaluating* their conjecture in *one or more* specific cases" (p.252). An inductive warrant uses a similar strategy to reduce uncertainty about the conclusion of an argument. Many examples of inductive warrants were used by participants in the current study. One example has already been discussed: Figure 4 shows an argument offered by Chris during his work on Conjecture 3. He quantitatively evaluated the conjecture for both small and large numbers, and thus felt that it should be true for all numbers.

In his response to Conjecture 4, having failed in his initial proof attempt, Andrew offered the following argument:

ANDREW: Let's make some experiments [LAUGHS]. OK, so the deficient numbers are, for example, 9. 9 is deficient. That's too big because, OK, 10 let's say, we've got 2, 5. Primes are apparently deficient.

INTERVIEWER: Primes are deficient?

ANDREW: Primes are always deficient, yeah, because the sum is equal to the number plus 1. Well, always [LAUGHS], no, or is it? no, even 2 is deficient, so it doesn't fail. Yeah, so apparently it works here. Yeah ok, so apparently, it seems to me that it's true.

INTERVIEWER: Why do you say that then? Because it works for 10?

ANDREW: Because [PAUSE] because, hmm. [LONG PAUSE]

At this point, after a long pause, Andrew began a proof attempt that eventually resulted in a correct proof. Andrew's argument is modelled in Figure 5. Based on one empirical evaluation, Andrew's uncertainty in the conjecture is sufficiently reduced to start a proof attempt. In the language of Balacheff (1988), Andrew conducted a 'crucial experiment' to persuade himself of the statement's probable truth.
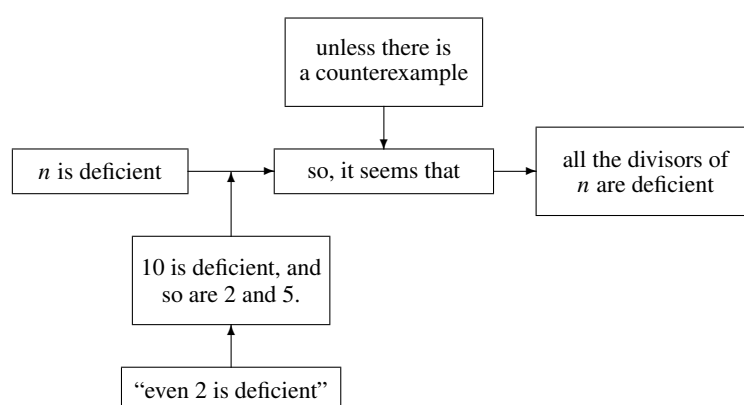

Figure 5. Part of Andrew's response to Conjecture 4.

In a similar example, David constructed a two stage argument to evaluate Conjecture 2. In the first stage he successfully showed that $\sigma(kn) \geq 2kn$ if n is perfect[3] (using the same argument deployed during his work on Conjecture 1), and during the second stage he tried to remove the possibility that kn is abundant:

DAVID: Why would it [THE EQUATION $\sigma(kn) \geq 2kn$] be a greater than? Umm, I don't know, why couldn't it be perfect? I mean you've got some possible counterexamples here, I mean, we might look for one of them, so does 6 divide into that? [LOOKS AT THE LIST OF EXAMPLES OF PERFECT NUMBERS] I don't know, no it doesn't does it? So does 6 divide into the next thing? So, I can't see any counterexamples there, and for example… So I guess, umm, what was we, what would I, umm, we need to find some divisors that aren't of the form $2m$ for m a divisor of n, don't we?

David's search for counterexamples is a failure, so he concludes that it is plausible that the statement is true, and attempts a proof. It is notable that before he

looked for possible counterexamples, David seemed unsure of whether it was true or not. The list of examples David used contained only 28, 496 and 8128; after noting that 6 did not divide any of these three perfect numbers, David was sufficiently satisfied that the conjecture was true (that is, his uncertainty was sufficiently reduced) that he began a proof attempt. The second stage of his argument is modelled in Figure 6.
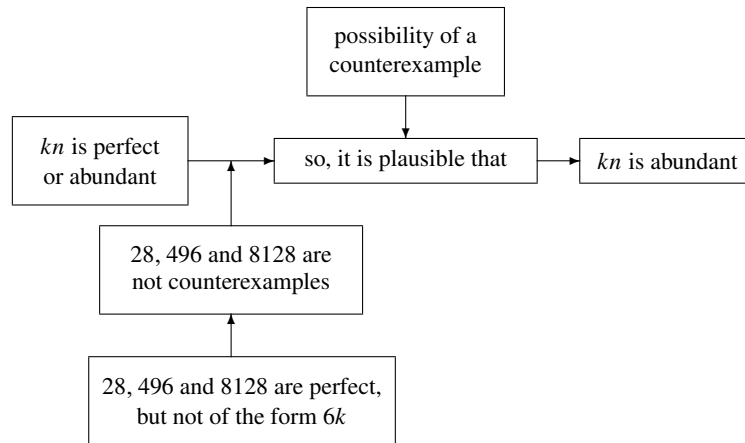

Figure 6. Part of David's response to Conjecture 2.

On many occasions, participant's degrees of belief in conclusions were affected by the use of examples. Two distinct strategies emerged: the use of examples (e.g. Figures 4 and 5) and the use of counterexamples (e.g. Figure 6). The first strategy involved the use of an example as a 'crucial experiment' (Balacheff, 1988) to test whether the Conjecture held in that case or not. If it did, a 'plausible' modal qualifier was used to link the conclusion with the data. The second strategy was somewhat different, in this case participants looked through a series of examples to see if they could find a counterexample. If they couldn't, again a 'plausible' modal qualifier was used. The second strategy seems closer to Balacheff's 'naive empiricism'.

In both cases, once participants sufficiently reduced their uncertainty in the conclusion, they attempted to prove it formally. No participant used an inductive warrant to attempt deduce *with certainty* that the conclusion followed from the data, and this would not have been expected, since they were all highly talented mathematicians. However, inductive warrants were widely used to reduce uncertainty in the conclusion, but they were moderated with appropriate modal qualifiers. Crucially, without incorporating the modal qualifier into the modelling of these arguments, an accurate picture of these types of reasoning would be impossible.

5. The structural-intuitive warrant-type.

The term 'structural-intuitive' is used to refer to a participant using observations about, or experiments with, some kind of mental structure, be it visual or otherwise, that persuades them of a conclusion. Often, but not necessarily always, this sort of reasoning appears to be of an intuitive type (in the sense of Fischbein, 1987).

Chris's reasoning about Conjecture 4 exemplifies the structural-intuitive warrant-type:

CHRIS: So if n is deficient then we get for free that umm, none of it's divisors are perfect, so every divisor must be deficient or abundant. Umm, it would seem odd if they were allowed to be deficient and abundant but not perfect. Because perfect is kind of the middle case, so it looks true.

This argument is modelled in Figure 7. Chris's warrant here is based on some intuitive understanding about how the properties of deficiency and abundancy should behave, and a realisation that if the conjecture was false, it would mean that these properties would have been broken.
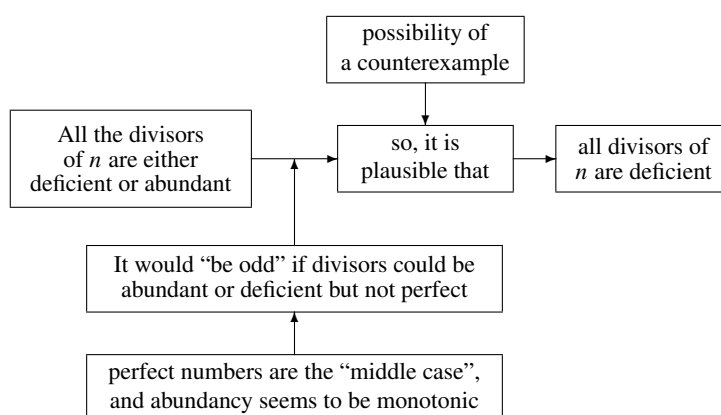

Figure 7. Part of Chris's response to Conjecture 4.

Conjectures 5 and 6 asked participants to consider the situation where m and n are abundant, and to decide whether they can conclude that $m + n$ or mn is abundant, respectively. These conjectures were particularly rich sources of structural-intuitive warrants. Conjecture 5 provoked many participants to argue that the conjecture was probably false. Take Fred, for example:

FRED: I think, going on instinct, it's probably false.

INTERVIEWER: Why?

FRED: Because, err, I mean, think of it... Another typical example, like, saying whether something is abundant is to do with it's divisors, so it's to do with things that

divide it, it's to do with multiples. And then, when you add two numbers to-
gether, it doesn't necessarily mean that any properties of the divisors stay the
same. I mean, like, I don't know, when you add 3 and 5. 3 and 5 have certain
divisors, but 8 has completely different divisors. Umm, but you never know.
So, but abundant is a very sort of wide statement, so, I mean, intuitively you'd
expect to apply to roughly half of all numbers, so maybe it's not so absurd to
think they would, err, that would hold. So I'll try.

Fred went on to try to prove the statement before abandoning his attempt and
looking for counterexamples.

Fred's behaviour here is interesting. Immediately after having read the
statement he seems sure that it is false, and justifies his intuition with a
structural-intuitive warrant based on the absence of a link between addition
and divisors. However, as he speaks, he appears to try to convince himself
not to trust his original intuition. Notwithstanding this initial change of mind,
Fred went on to discover that the statement is indeed false, and realised that
looking for a counterexample is straightforward.

Similar structural-intuitive warrants were used by other participants:

CHRIS: Right, so if m and n are abundant, then $m + n$ is abundant. That doesn't look
true.

INTERVIEWER: Why not?

CHRIS: Because the factors of $n + m$ don't really have anything to do with the factors
of n or m. So it should be fairly easy to construct a counterexample. I say that
[LAUGHS]. So if I pick two nice abundants, umm. . .

Chris went on to find a counterexample.

It is interesting to compare participants' immediate responses to Conjec-
ture 5 with their corresponding responses to Conjecture 6. Here is Chris's
response:

CHRIS: [READS CARD] Right, so if n and m are abundant then nm is abundant. That
looks more plausible, because they're going to share factors.

This excerpt from Chris illustrates the use of similar structural-intuitive war-
rants to that used in Conjecture 5, but here it carries plausibility rather than
implausibility. In the case of both conjectures the warrants act to reduce
uncertainty sufficiently to continue the argument in the given direction.

The pattern of responses to Conjecture 5 was uniform; participants' initial
intuitions gave them a structural-intuitive warrant which they used to decide
that the conjecture was unlikely to be true. This then directed their attempts
at refuting the conjecture. Since they had deduced that the conclusion was
unlikely to be true, looking for a counterexample was the most appropri-
ate strategy. In Conjecture 6 the situation was reversed. Participants used

structural-intuitive warrants to determine that the conjecture was likely to be true, and then based their decision to look for a proof on this judgement.

The reliability of intuition in mathematics has been a recurring subject of discussion by mathematicians and philosophers. Hahn, in a famous essay written in 1933 (reprinted as Hahn, 1960), even went as far as to argue that intuition is entirely unreliable and should be "expelled" from mathematical reasoning. To back up his position Hahn gave several examples of counter-intuitive 'monsters': a map of three regions which meet each other at every point along one border, and a curve which intersects itself at every point (e.g. Moore, 1900; Whyburn, 1942). These objects, Hahn argued, are impossible to reconcile with intuition, and thus he argued that intuition needs to be removed from all mathematical reasoning. Other authors have disagreed with this analysis, pointing out that although intuition may sometimes be misleading it is essential for giving direction to mathematical research (Feferman, 2000; Poincaré, 1905). The data from this study support this latter stance. Participants used their intuitive structures to establish a belief in whether the conclusion follows from the data. Structural-intuitive warrants were widely used to reduce uncertainty.

As we have argued throughout this paper, the full Toulmin scheme allows us to discuss the wider issue of reducing uncertainty, while the restricted version (without modal qualification or with the trivial modal qualification of 'in all cases') focuses narrowly on preserving certainty.

Structural-intuitive warrants, whilst reducing uncertainty, do sometimes support incorrect conclusions. In several of the interviews the issue of whether abundant numbers need to be even developed naturally. When Fred was working on Conjecture 5, for example, he considered what sorts of numbers were likely to be abundant:

FRED: Err, well my thinking is, odd numbers are not [abundant], generally because. . .

INTERVIEWER: When you say "generally" what do you mean?

FRED: Just the general idea, because, like, if a number is even then one of its divisors is half the number, which is a pretty big chunk, but if a number is odd it's missing a big chunk.

INTERVIEWER: So, you reckon no odd numbers are abundant?

FRED: I think that's quite unlikely.

During the course of his interview Ben was asked directly whether he thought all abundant numbers were even, his argument was similar to Fred's:

INTERVIEWER: I mean that's an incredibly difficult question, but what's your sort of, if you had to have a stab in the dark about it?

BEN: Umm, [LONG PAUSE] I think it might have to be true.

INTERVIEWER: Why?

BEN: [LONG PAUSE] I think if they're odd, you lose too much of the, sort of, sequence that you can't divide into, if you get what I mean, because you can't, you can't divide past, so say it was odd and the first one was 3, you'd only have ones up to $\frac{n}{3}$ and then n, whereas if you go up to $\frac{n}{2}$ you get a lot more, well in theory, you could get a lot more possible divisors, so it's based on a sort of size argument rather than anything particularly... But, intuitively, even numbers would certainly be more likely to be abundant than odd numbers.

The structure of Ben and Fred's arguments is modelled in Figure 8. Both used structural-intuitive warrants about the properties of abundant numbers, using understanding that they had built up through working on previous conjectures.
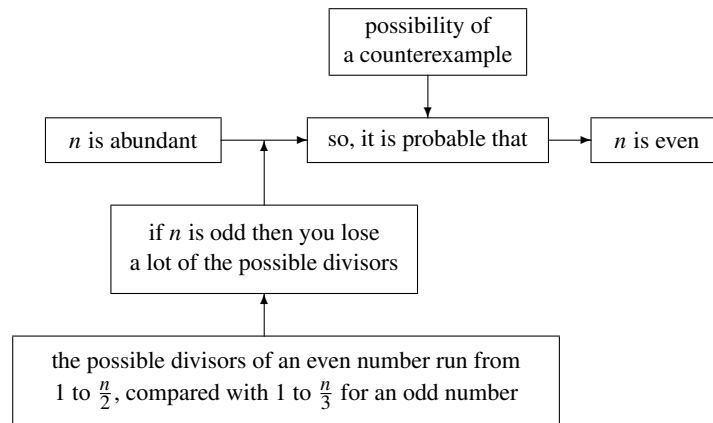

Figure 8. Ben and Fred's argument regarding the parity of abundant numbers.

Notwithstanding Ben and Fred's structural-intuitive warrants, there are in fact an infinite number of odd abundants, with 945 the first. Indeed, surprisingly, it is possible to construct an abundant number whose smallest divisor is arbitrarily high. So Ben and Fred's (apparently reasonable) structural-intuitive warrant that odd abundants are unlikely, as the divisors can only 'live' in, at best, the lowest third of the number turns out to be wrong. Abundants can be constructed where the divisors live in the lowest kth of the number where k is as high as you like. Thus the structural-intuitive warrants in this case considerably reduced their uncertainty in the conjecture that every abundant number is even, despite this being false.

6. The deductive warrant-type.

Harel (2001) referred to the most sophisticated proof scheme as the 'deductive-modern-axiomatic' scheme: people who have this scheme use deductions

from axioms to establish truth. A similar notion is the basis of the deductive warrant-type: formal mathematical justifications are used to warrant the conclusion of the argument in question. These justifications can be of various sorts: deductions from axioms, algebraic manipulations, or the use of counterexamples would all be classified as deductive warrants.

For professional mathematicians, a deductive warrant is seen as carrying formal mathematical necessity: an argument that uses a deductive warrant admits no effective rebuttal. It could be argued that, in complex proofs, mathematicians *do* sometimes have non-trivial qualifiers and rebuttals – such as 'unless there is a flaw in my argument' – but the aim of these forms of argument is to minimise this. Thus, while for professional mathematicians the inductive and structural-intuitive warrant types *aim* to reduce uncertainty, the deductive warrant *aims* to *remove* uncertainty. Although this is the case for professional mathematicians, it may not be for all students: the potential for constructing an inappropriate matching between deductive warrants and modal qualifiers is discussed later in the paper.

Examples of deductive warrants were present throughout our data. For example, when Andrew was working on Conjecture 2 he produced the following argument (having used a different approach in Conjecture 1):

ANDREW: OK, so if n is perfect, then kn is abundant, for any k. OK, so what does it, yeah it looks, so what does it mean? Yeah so if n is perfect, and I take any p_i which divides this n, then afterwards the sum of these p_is is $2n$. This is the definition. Yeah, ok, so actually we take kn, then obviously all kp_i divide kn, actually, we sum these and we get $2kn$. Plus, we've got also, for example, we've also got k dividing this, dividing kn. So we need to add this. As far, as basically, there is no disquiet, k would be the same as this. Yeah. And, how would this one go? [LONG PAUSE]

INTERVIEWER: So we've got the same problem as up here [CONJECTURE 1] but in general? With a. . . ?

ANDREW: Yeah. Umm, can we find one? Right, so I don't know. Some example.

INTERVIEWER: I've got some examples for you.

ANDREW: You've got examples of some perfect numbers? OK, so 12, we've got $1 + 2 + 3 + 4 + 6$, then, ok, $+12$. [MUTTERS] But this is not? OK, perfect, I wanted perfect numbers. OK, so let's say 6. Yeah, and we've got $1 + 2 + 3 + 6$ and actually we take 2×6 which is 12. Then yes, I've got divisors 2, 4, 6, 12. Plus I claim we've got also divisors. Yeah! actually it's simple because, err, because err, the argument is that we've also got 1 which is divisor, and this divisor is no longer contained here if we multiply.

There are two quite distinct stages to this argument, both use a deductive warrant. The first establishes that n must be either abundant or perfect. At this stage Andrew realises that he needs to reject the perfect case, and uses the generic example of 2×6 to do this (Balacheff, 1988). In both stages

the conclusion follows necessarily from the data. Although Andrew believed that no rebuttals were possible, in fact there is a possible rebuttal, the case where $k = 1$. When prompted by the interviewer, he immediately recognised this trivial case, and modified the conclusion appropriately. Note, again, that though formally Andrew's conclusion was incorrect until he made this modification, his deductive warrant had effectively removed his uncertainty.

Not all deductive warrants are of this form, where conclusions are deduced from data by logical implications. Sometimes, for example, participants used counterexamples to warrant their conclusions. Here is part of Edward's argument in Conjecture 1. After he had successfully shown that all *proper* multiples of 6 are abundant, he turned his attention to 6 itself:

EDWARD: Yeah. [. . .] so this leaves the special case of $n = 6$. . . which is a perfect number, I just know that. Or is it? Hmm. $1 + 2 + 3 + 6 = 12$, yeah, so it's a perfect number. Abundant, does that mean greater or equal to, or just greater? So, that's, that's a counterexample, so when $n = 6$ it's not an abundant number. So. . . we've got rid of the 'if'.

The key distinction, then, in the manner in which participants used deductive warrants compared to those discussed in previous sections, is that deductive warrants were matched with absolute modal qualifiers and that no rebuttals were admitted. They are the types of warrants discussed by Aberdein (2005) during his analysis of formal mathematical proofs. However, as we demonstrate in the next section, this appropriate matching between deductive warrants and absolute modal qualifiers may not always happen.

7. Discussion.

In the previous three sections three different warrant-types have been discussed, with examples of the use of all three by highly talented mathematics research students. We do not claim that this threefold classification is exhaustive: it is possible that the type of problems used in this research restricted the range of warrant-types we were likely to uncover; putting the participants in a situation which modelled mathematical research processes more realistically, or choosing problems from a different area, might have led to us seeing different types of argumentation being used. Indeed, we do not even argue that a warrant may belong to a distinct warrant-type. For example, Chris's argument, given in Figure 4, seems to combine both inductive properties (he quantitatively evaluated several examples) with structural-intuitive properties (he believed that number theory tends to be monotonic).

The restricted form of Toulmin's (1958) scheme used by earlier researchers to model mathematical argumentation constrains us to think only in terms of

arguments with absolute conclusions. In our data we see talented mathematicians using a much wider range of argumentation; a range which needs the full Toulmin scheme to model. The key, however, is that these wider forms of argument are not put forward by the participants as proofs: the inductive and structural-intuitive warrants they use are paired with entirely appropriate modal qualifiers which indicate they appreciate they have *reduced but not removed* their uncertainty in the conjectures. A proof stage (which can generally be modelled with empty or trivial qualifiers and rebuttals) appeared only later in the development of their arguments.

7.1. MATHEMATICAL DEVELOPMENT.

When developing their proof schemes framework, Harel and Sowder (1998) argued that in order to succeed at advanced level mathematics students must replace inductive, transformational and external proof schemes with a deductive scheme:

> "[T]he goal of instruction must be unambiguous; namely, to gradually refine current students' proof schemes toward the proof scheme shared and practiced by the mathematicians of today" (Harel, 2001)

Similar arguments have been made by other researchers. Tall (2004), for example, argued that as students deepen their cognitive development, their 'warrants for truth' (in the sense of Rodd, 2000) also deepen, hopefully with the result that formal proof becomes the only acceptable warrant. Tall referred to this process as moving through the 'three worlds' of mathematics: from the embodied and proceptual worlds through to the formal-axiomatic world (Tall, 2004).

We would certainly agree with Harel (2001) and Tall (2004) when they suggest that no student will be successful at advanced mathematics if they accept a conclusion *with certainty* on the basis of non-deductive warrants. However, we argue that our data indicates these non-deductive warrant-types play a crucial role in mathematical argumentation, *as long as they are paired with appropriate modal qualifiers*. We suggest that, when a person enters Tall's (2004) formal-axiomatic world, or when they develop Harel and Sowder's (1998) axiomatic-deductive proof scheme, rather than reducing the range of warrant-types they use, they retain the use of the warrants that have been used in previous 'worlds' or 'proof schemes', but *they qualify them appropriately* (where appropriateness is defined by expert practice). Our data show that mathematicians do not abandon inductive and intuitive arguments; instead, they learn to pair them with appropriate modal qualifiers and rebuttals. It is this pairing that is so crucial to successfully developing as a mathematician.

7.2. Inappropriate warrant-qualifier pairings.

There is evidence that this key skill – the ability to appropriately pair warrants with modal qualifiers – is not always present. In this section we briefly discuss two examples taken from the literature which show students constructing inappropriate pairings.

Weber (2003) reported a student's purported proof of the statement "for every odd integer n, $n^2 - 1$ is divisible by 8":

> "$1^2 - 1 = 0$ which is divisible by 8. $3^2 - 1 = 8$ which is divisible by 8. $5^2 - 1 = 24$ which is divisible by 8. And so on. Therefore if n is odd, $n^2 - 1$ is divisible by 8."

As a consequence of their use of an inductive warrant, Harel and Sowder (1998) would describe this student as having an inductive proof scheme, and in terms of Tall's (2004) framework they are yet to reach the axiomatic-formal world. In terms of our framework for modelling mathematical argumentation, this argument has a modal qualifier which is inappropriately matched with its warrant. The key difference between our perspective and that of Harel and Sowder and Tall, however, is our examination of the ways in which non-deductive warrants may be used appropriately in mathematics. The use of an inductive warrant is not inappropriate per se, it is only when it is inappropriately paired with an absolute modal qualifier that the argument becomes problematic.

In this case, Weber's (2003) student inappropriately paired a non-deductive warrant with an absolute qualifier, but there are also reported cases of students doing the reverse: pairing deductive warrants with non-absolute qualifiers.

As part of his work on learning styles, Simpson (1995) discussed responses to the so-called 'Arithmagons' problem (Mason et al., 1982):

> "A secret number is assigned to each vertex of a triangle. On each side of the triangle is written the sum of the secret numbers at its ends. Find a simple rule revealing the secret numbers."

Simpson reported one student's behaviour:

> "Having been asked to prove a result which she had stated after some time working on the [Arithmagons problem] she wrote a quite delightful little proof which, though just essentially algebraic manipulation, made me feel that she had grasped the essence of the problem and gave a quite general solution.
>
> On the next page, she wrote 'I wonder if it works for big numbers?' " (Simpson, 1995; see also Duffin and Simpson, 1993).

This student, despite having presented an apparently perfect deductive proof, did not pair it with an absolute modal qualifier. For her, the deductive warrant she had written only allowed her to conclude that the statement was true

about small numbers. The possibility that large numbers could form a rebuttal remained a concern for her.

In short, this student used a deductive warrant successfully, but was unable to qualify it suitably. Similar examples have been reported by Fischbein (1982). Such examples indicate that the use of deductive warrants alone is not sufficient in advanced mathematics: they must be paired with appropriate modal qualifiers.

8. Final remarks.

Earlier researchers have attempted to model mathematical arguments using a restricted version of Toulmin's (1958) argumentation scheme, in which the role of the modal qualifier and rebuttal were downplayed, marginalised or omitted entirely. In this paper we have argued that this restricted version is inadequate for accurately modelling the full range of arguments constructed by mathematics students. Our data, collected from highly successful postgraduate mathematics students, shows frequent use of non-deductive warrants to deduce non-absolute conclusions and highlights that these forms of argument are crucial for them in the process of solving the problem. This type of argumentation would be impossible to accurately model using the restricted scheme advocated by earlier researchers.

Furthermore, we have argued that omitting the role of the modal qualifier in models of mathematical arguments constrains us to consider only arguments with absolute conclusions, and, consequently, to undervalue non-deductive warrants in advanced mathematics. Whereas the various non-deductive *proof schemes* described by Harel and Sowder (1998) certainly have no place in advanced mathematics classes, we suggest the equivalent non-deductive *warrant-types* do have an important role to play in mathematical argumentation. While the purpose of this paper was not intended to directly affect pedagogy, examining the nature of the arguments put forward here suggests a change: perhaps the goal of instruction should not be to remove any trace of inductive or intuitive reasoning from students' arguments, but to ensure that they qualify these sorts of warrants appropriately.

Appendix

Conjecture 1. A number is abundant iff it is a multiple of 6. *False.* 20 is a counterexample to the 'only if' statement, it is abundant, but not a multiple of 6. 6 is a counterexample to the 'if' statement, it is a multiple of 6 and is perfect, not abundant. However, all multiples of 6 apart from 6 are abundant. Take $n = 6k$ for some $k \neq 1$. Then n has, at least, $1, k, 2k, 3k, 6k$ as distinct divisors which sum to greater than $12k = 2n$.

Conjecture 2. If n is perfect, then kn is abundant for any $k \in \mathbb{N}$. *False.* If $k = 1$ then kn is clearly not abundant. The statement is true provided $k \neq 1$. Suppose n is perfect with divisors $d_1, d_2, \ldots, d_r$ (i.e. $2n = d_1 + \ldots + d_r$). Then kn has amongst its divisors $1, kd_1, kd_2, \ldots, kd_r$, and these sum to $2kn + 1 > 2kn$. So kn is abundant.

Conjecture 3. If p_1 and p_2 are primes, then $p_1 p_2$ is not abundant. *True.* If p_1, p_2 are distinct, we need to show that $1 + p_1 + p_2 + p_1 p_2 \leq 2p_1 p_2$. This can be rearranged to $(p_1 - 1)(p_2 - 1) \geq 2$ which is clearly true as at least one of p_1, p_2 is greater than 2. If $p_1 = p_2$, the inequality reduces to $1 + p_1 \leq p_1^2$ which is true for all integers greater than 1.

Conjecture 4. If n is deficient, then every divisor of n is deficient. *True.* Consider the contrapositive: if n is not deficient, then kn is not deficient. Suppose n has divisors $d_1, \ldots, d_r$ and that $d_1 + \ldots + d_r \geq 2n$. Then the set of divisors of kn contains $1, kd_1, kd_2, \ldots, kd_r$. And we know that $1 + kd_1 + \ldots + kd_r \geq 2kn + 1$. Therefore kn is not deficient.

Conjecture 5. If n and m are abundant, then $n + m$ is abundant. *False.* 20 and 12 are abundant, but 32 is deficient ($1 + 2 + 4 + 8 + 16 + 32 = 63 < 2 \times 32$).

Conjecture 6. If n and m are abundant, then nm is abundant. *True.* This is a specialisation of the claim that all multiples of abundant numbers are abundant, so it is true.

Conjecture 7. If n is abundant, then n is not of the form p^m for some natural m and prime p. *True.* Consider the contrapositive: if n is of the form p^m then n is not abundant. The divisors of p^m are $1, p, p^2, \ldots, p^m$, so we need to show that $1 + p + \ldots + p^m \leq 2p^m$. Trivially $p^m(p-1) \geq p^m - 1$. But $p^m - 1 = (p^{m-1} + p^{m-2} + \ldots + p + 1)(p-1)$, so $p^m \geq p^{m-1} + p^{m-2} + \ldots + p + 1$ and hence $2p^m \geq p^m + p^{m-1} + p^{m-2} + \ldots + p + 1$.

Acknowledgements

Notes

[1] Toulmin's (1958) use of the word 'warrant' is not identical to how the term has been used by some of the mathematics education literature. In the terms we will use in this paper, Rodd (2000) saw a warrant as *removing uncertainty*, whereas Toulmin was more flexible, accepting that a warrant can be qualified with a modal qualifier to *reduce uncertainty*.

[2] There has been some discussion in the mathematics education literature regarding the differences and similarities between 'argumentation' and 'proof' (e.g. Boero, 1999, Duval, 1991). Adopting the standard terminology used by Toulmin (1958), we, and presumably other researchers who employ his scheme, interpret mathematical proof to be a type of argument.

[3] Here $\sigma(n)$ denotes the sum of the divisors of n.

References

Aberdein, A.: 2005, 'The uses of argument in mathematics'. *Argumentation* **19**, 287–301.

Aberdein, A.: 2006, 'The informal logic of mathematical proof'. In: R. Hersh (ed.): *18 Unconventional Essays on the Nature of Mathematics*. New York: Springer, pp. 56–70.

Alcolea Banegas, J.: 1998, 'L'Argumentació en Matemàtiques'. In: E. C. i Moya (ed.): *XIIè Congrés Valenciaà de Filosofia*. Valencià, pp. 135–147, Diputació de Valencià.

Balacheff, N.: 1988, 'Aspects of Proof in Pupils' Practice of School Mathematics'. In: D. Pimm (ed.): *Mathematics, Teachers and Children*. London: Hodder & Stoughton, pp. 216–235.

Boero, P.: 1999, 'Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education'. International Newsletter on the Teaching and Learning of Mathematical Proof, July/August 1999.

Bromley, D. B.: 1986, *The case-study method in psychology and related disciplines*. Chichester: John Wiley & Sons.

Burton, L.: 2004, *Mathematicians as Enquirers: Learning about Learning Mathematics*. Dordrecht: Kluwer.

Duffin, J. and A. Simpson: 1993, 'Natural, Conflicting and Alien'. *Journal of Mathematical Behavior* **12**, 313–328.

Duval, R.: 1991, 'Structure du raisonnement déductif et apprentissage de la démonstration'. *Educational Studies in Mathematics* **22**, 233–261.

Evens, H. and J. Houssart: 2004, 'Categorizing pupils' written answers to a mathematics test question: 'I know but I can't explain". *Educational Research* **46**, 269–282.

Feferman, S.: 2000, 'Mathematical Intuition vs. Mathematical Monsters'. *Synthèse* **125**, 317–332.

Fischbein, E.: 1982, 'Intuition and proof'. *For the Learning of Mathematics* **3**(2), 9–18.

Fischbein, E.: 1987, *Intuition in Science and Mathematics*. Dordrecht: Reidel.

Ginsburg, H.: 1981, 'The Clinical Interview in Psychological Research on Mathematical Thinking: Aims, Rationales, Techniques'. *For the Learning of Mathematics* **1**(1), 4–11.

Hadamard, J.: 1945, *The Psychology of Invention in the Mathematical Field*. New York: Dover Publications, 1954 edition.

Hahn, H.: 1933/1960, 'The crisis in intuition'. In: J. R. Newman (ed.): *The World of Mathematics*, Vol. 3. London: Allen and Unwin, pp. 1956–1976.

Harel, G.: 2001, 'The development of mathematical induction as a proof scheme: A model for DNR-based instruction'. In: S. Campbell and R. Zazkis (eds.): *Learning and teaching number theory*. Westport, CT: Ablex Publishing Corp, pp. 185–212.

Harel, G. and L. Sowder: 1998, 'Students' Proof Schemes: Results from Exploratory Studies'. In: A. H. Schoenfeld, J. Kaput, and E. Dubinsky (eds.): *Research in Collegiate Mathematics III*. Providence, RI: American Mathematical Society, pp. 234–282.

Hoyles, C. and D. Küchemann: 2002, 'Students' Understanding of Logical Implication'. *Educational Studies in Mathematics* **51**(3), 193–223.

Knipping, C.: 2003, 'Argumentation structures in classroom proving situations'. In: M. A. Mariotti (ed.): *Proceedings of the Third Congress of the European Society for Research in Mathematics Education*. Bellaria, Italy, ERME.

Krummheuer, G.: 1995, 'The ethnology of argumentation'. In: P. Cobb and H. Bauersfeld (eds.): *The Emergence of Mathematical Meaning: Interaction in Classroom Cultures*. Hillsdale: Erlbaum, pp. 229–269.

Mariotti, M. A.: 2006, 'Proof and proving in mathematics education'. In: A. Gutiérrez and P. Boero (eds.): *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future*. Rotterdam: Sense, pp. 173–204.

Markowitz, L. M. and R. D. Tweney: 1981, 'Confirmatory and Disconfirmatory Heuristics in Mathematical Reasoning'. Unpublished manuscript.

Mason, J., L. Burton, and K. Stacey: 1982, *Thinking Mathematically*. London: Addison-Wesley.

Moore, E. H.: 1900, 'On Certain Crinkly Curves'. *Transactions of the American Mathematical Society* **1**, 72–90.

Pedemonte, B.: 2005, 'Quelques outils pour l'analyse cognitive du rapport entre argumentation et démonstration'. *Recherches en Didactique des Mathématiques* **25**, 313–348.

Pedemonte, B.: in press, 'How can the relationship Between Argumentation and Proof be analysed?'. To appear in *Educational Studies in Mathematics*.

Poincaré, H.: 1905, *Science and Hypothesis*. London: Walter Scott Publishing.

Rodd, M. M.: 2000, 'On Mathematical Warrants: Proof Does Not Always Warrant, and a Warrant May Be Other Than a Proof'. *Mathematical Thinking and Learning* **2**, 221–244.

Simosi, M.: 2003, 'Using Toulmin's Framework for the Analysis of Everyday Argumentation: Some Methodological Considerations'. *Argumentation* **17**, 185–202.

Simpson, A.: 1995, 'Focusing on Student Attitudes to Proof'. *Teaching and Learning Undergraduate Mathematics Newsletter* **3**.

Tall, D. O.: 2004, 'Building theories: The three worlds of mathematics: A comment on Inglis'. *For the Learning of Mathematics* **23**(3), 29–32.

Thurston, W. P.: 1994, 'On proof and progress in mathematics'. *Bulletin of the American Mathematical Society* **30**, 161–177.

Toulmin, S.: 1958, *The uses of argument*. Cambridge: CUP.

Toulmin, S.: 2001, *Return to reason*. Cambridge, MA: Harvard University Press.

Toulmin, S., R. Rieke, and A. Janik: 1984, *An introduction to reasoning*. New York: Macmillan, second edition.

Weber, K.: 2003, 'Students' Difficulties with Proof'. MAA Research Sampler 8, available at www.maa.org.

Weber, K. and L. Alcock: 2005, 'Using warranted implications to understand and validate proofs'. *For the Learning of Mathematics* **25**(1), 34–38.

Whyburn, G.: 1942, 'What is a Curve?'. *American Mathematical Monthly* **49**, 493–497.

Yackel, E.: 2001, 'Explanation, justification and argumentation in mathematics classrooms'. In: M. van den Heuvel-Panhuizen (ed.): *Proceedings of the 25th International Conference on the Psychology of Mathematics Education*, Vol. 1. Utrecht, Holland, pp. 9–23, IGPME.

Address for Offprints:
Matthew Inglis
Learning Sciences Research Institute
University of Nottingham
Jubilee Campus
Wollaton Road
Nottingham, UK.
e-mail: matthew.inglis@nottingham.ac.uk