

Author's name: **Andrew Aberdein**

Title: **Mathematical wit and mathematical cognition**

Author's affiliation: Department of Humanities & Communication, Florida Institute of
Technology

Author's phone: +44 1704 874971

Author's email: aberdein@fit.edu

Suggested running head: Mathematical wit and mathematical cognition

Keywords: argumentation schemes; citation; generalization; humour; intuition;
mathematical practice; mathematical reasoning; retrodution.

MATHEMATICAL WIT AND MATHEMATICAL COGNITION

ANDREW ABERDEIN

ABSTRACT. The published works of scientists often conceal the cognitive processes that led to their results. Scholars of mathematical practice must therefore seek out less obvious sources. This paper analyzes a widely circulated mathematical joke, comprising a list of spurious proof types. An account is proposed in terms of argumentation schemes: stereotypical patterns of reasoning, which may be accompanied by critical questions itemizing possible lines of defeat. It is argued that humour is associated with risky forms of inference which are essential to creative mathematics. The components of the joke are explicated by argumentation schemes devised for application to topic neutral reasoning. These in turn are classified under seven headings: retroduction; citation; intuition; meta-argument; closure; generalization; and definition. Finally, the wider significance of this account for the cognitive science of mathematics is discussed.

1. INTRODUCTION

1.1. *The joke in question*

The role of humour in mathematical cognition has seldom been addressed. This paper will focus on a single mathematical joke,¹ albeit one with multiple parts. It is a list of spurious proof methods which has circulated on mathematical bulletin boards, real and virtual, since the 1980s. It deserves attention for several reasons. Firstly, it is well known among mathematicians. Secondly, the joke is concerned with a central aspect of mathematical practice: proof, and especially borderline cases of proof. As Paul Renteln and Alan Dundes observe, the list

hints at the anxieties felt by many mathematicians regarding the degree to which mathematical truth is dependent upon the trustworthiness of previous results. This anxiety is exacerbated by the fact that some mathematicians have a less rigorous proof style than other mathematicians (Renteln & Dundes, 2005, p. 28).

Thirdly, its dissemination is sufficiently documented to observe that the list has remained fairly stable over time. Although some items incorporate older material, 24 of them were compiled as (Angluin, 1983). A rearranged version, with three additions and no attribution, may be found in (Cherkaev & Cherkaev, nd).² This website, which contains many other mathematical jokes, is acknowledged as a source by Renteln and Dundes (2005, p. 24), who reproduce the list in the Cherkaevs' rearrangement, omitting their most recent addition. Although the list is obviously not exhaustive of problematic forms of proof, the survival of most of the items through these changes and the comparative unoriginality of the new items suggest that it is broadly comprehensive. Lastly, as Homer Simpson would say, 'It's funny because it's true.' Many of the

¹It lacks some of the structural features common to many jokes, but it does meet an informal definition: 'a relatively short text which, for a given cultural group, is recognizable as having, as its primary purpose, the production of an amused reaction in its reader/hearer, and which is typically repeatable in a wide range of contexts' (Ritchie, 2004, p. 15). However, nothing that follows rests on whether the text in question is a joke or some other sort of wit.

²The Cherkaevs assert, with more candour than legal acumen, that 'To our mind, a joke goes to "public domain" immediately after being created or modified and there should be no authorship in it.' A search on archive.org reveals the list was on their site by 2000, and was added to in 2003 and 2004.

items are only lightly exaggerated, and many similar examples of actual mathematical practice could be given, although for reasons of space only a few are documented here.

1.2. *Humour theory*

Some surveys distinguish as many as eight conflicting categories of humour theory; it seems correct to say that each category succeeds in accounting for an important feature, but typically at the expense of other important features (Hurley, Dennett & Adams, 2011, p. 37). Amongst the most influential are superiority, release, and incongruity theories, each of which has distinguished champions. On the superiority account, associated with Hobbes, laughter proceeds from the recognition of the joke teller's superiority over the butt of the joke. On the release account, associated with Freud, laughter is a release of tension resulting from the overcoming of an inhibition. On the incongruity account, associated with Kant, laughter results from the resolution of an apparent absurdity.

Humour has been the subject of empirical inquiry in all the major domains of psychological research (Martin, 2007, p. 29). In social psychology, it has been shown how humour can be used to socialize members of a group into the acceptance of norms important to the group without the hostility that may result from more direct criticism (Martin, 2007, p. 120). As Gary Fine and Michaela De Soucey conclude from their studies of mushroom collectors and meteorologists, 'Joking themes permit the smoothing of conversation, sharing traditions in the name of cohesion, separating the group from others, and securing the conformity of members' (2005, p. 17). All of these activities are just as important to the acculturation of student mathematicians, a central activity of the communities in which Angluin's list has circulated. So it seems reasonable to conjecture that the items on the list reflect norms that mathematicians consider important.

Humour is more intimately related to cognition: jokes may be seen as byproducts of defeasible cognitive heuristics. For example, Marvin Minsky has argued that 'Common sense logic is too unreliable for practical use. It cannot be repaired, so we must learn to avoid its most common malfunctions. Humor plays a special role in learning and communicating about such matters. . . . Its forms include graciously disarming ways to instruct others about inappropriate behavior and faulty reasoning' (1983, pp. 172 f.). A more sophisticated articulation of this insight characterizes humour as a reward mechanism for 'maintaining epistemic data-integrity' (Hurley et al., 2011, p. 127). On this account, humour has arisen through natural selection and

cultural evolution as compensation for the slapdash nature of much cognition. Our minds prefer fast and unreliable heuristics over reliable but resource-intensive alternatives. Left unchecked, this process would leave us with increasingly contradictory belief sets, leading to erratic and risky behaviour. Hence, there is selective pressure for a mechanism which will incentivize us to reconcile these incongruities. We experience the incentive as mirth.

1.3. *Fallacy theory*

Some of the items on Angluin’s list resemble fallacies, so it might be supposed that the analysis of the list could be subsumed under a general account of (mathematical) fallacy. However, fallacy theory is a much less developed field than humour theory. After centuries of stagnation, there has been significant progress in the study of informal fallacies (beginning with Hamblin, 1970), but this work has only begun to be applied to mathematical fallacies (see, for example, Krabbe, 2008; Aberdein, 2010). Most treatments of mathematical fallacies exhibit three significant shortcomings. Firstly, they emphasize what one may, following Bacon, call ‘gross fallacies’—cases where it is obvious that something is wrong and the problem is (only) to determine what—rather than ‘subtle fallacies’ which superficially appear sound (for this distinction, see Aberdein, 2010, p. 273). Secondly, surveys of mathematical fallacy are seldom more than lists, perhaps subdivided by the field or difficulty of the mathematics involved (for example, Maxwell, 1959; Barbeau, 2000). One exception does offer a ‘classification of exercises in refuting false mathematical arguments’ (Bradis, Minkovskii & Kharcheva, 1959, pp. 6 ff.). Some of the entries, such as ‘Extension to exceptional cases’ correspond to aspects of reasoning discussed in this paper, but many of them are by definition always erroneous. This leads to the third shortcoming of such surveys, that the emphasis is placed exclusively on irredeemably bad arguments, without exploring their interdependence with good (but not fully rigorous) reasoning.

1.4. *Argumentation schemes*

The account proposed in this paper utilizes *argumentation schemes*, stereotypical patterns of reasoning. They have an ancient origin, but have attracted renewed interest in recent years as a device for the analysis and evaluation not just of fallacies but of non-deductive argumentation in general. Deductive inference rules may be thought of as special cases of argumentation schemes, but the novelty of the approach rests in the application of defeasible schemes to informal reasoning. The widest survey to date distinguishes more than

90 schemes (Walton, Reed & Macagno, 2008, pp. 308 ff.). Many of these are specialized instances of a general scheme:

Argumentation Scheme 1.1 (Defeasible Modus Ponens).

As a rule, if P , then Q .

P .

It is not the case that there is an exception to the rule that if P , then Q .

Therefore, Q . (Walton et al., 2008, p. 366)

The defeasible nature of the reasoning is captured by an additional device, *critical questions*, which itemize possible lines of response. For the use of a scheme to be evaluated positively, its proponent must make adequate replies to such of these questions as may be raised by a respondent. We may distinguish different degrees of success, that is degrees of confidence with which the conclusion should be accepted, by requiring more or less extensive answers.

Although the argumentation scheme methodology was developed for non-mathematical contexts, it is applicable to mathematical reasoning in two ways: firstly, by using topic neutral schemes; and secondly, by developing specifically mathematical schemes. This paper pursues the first approach.³ The classification of argumentation schemes is an open problem, complicated, perhaps intractably, by the multiple dimensions of similarity between schemes. Walton proposes a threefold division between reasoning, source-based arguments, and the application of rules (Walton et al., 2008, pp. 348 f.). A more fine-grained version of this classification is employed below to highlight problematic aspects of mathematical argumentation.

2. RETRODUCTION

Much mathematical reasoning is sound deductive inference, secure from the anxieties that mathematical jokes reflect. But not all of it. One of the most ancient techniques of mathematical problem-solving is retroduction, or working backwards. In one of its aspects this was known to ancient mathematicians as analysis, from ‘ἀνάπαλιν λύσιν’ or ‘solution backwards’ (Heath, 1908, p. xlvi). (The term has, of course, since acquired other senses.) The steps which the analysis traces should be deductively sound, and once the

³For more on the first approach, see (Aberdein, 2010; Dove, 2009); for the second approach, see (Pease & Aberdein, 2011).

process is complete the order of the steps may be reversed to yield a conventional deductive proof. This is familiar in the strategy of reducing an open problem to one that has already been solved. Of course, the target must be chosen carefully, as one of Angluin's spurious proofs indicates:

Proof by reduction to the wrong problem: 'To see that infinite-dimensional colored cycle striping is decidable, we reduce it to the halting problem.'⁴ (Angluin, 1983, p. 17)

Multi-step argumentation is characteristic of, but not unique to mathematical reasoning. It may be understood in terms of the following scheme:

Argumentation Scheme 2.1 (Argument from Gradualism).

Premise: Proposition A is true (acceptable to the respondent).

Premise: There is an intervening sequence of propositions, $B_1, B_2, \dots, B_{n-1}, B_n, C$ such that the following conditionals are true: If A then B_1 ; if B_1 then B_2 ; \dots ; if B_{n-1} then B_n ; if B_n then C .

Premise: The conditional 'If A then C ' is not, by itself, acceptable to the respondent, nor are shorter sequences from A to C (than the one specified in the second premise) acceptable to the respondent.

Conclusion: Therefore, the proposition C is true (acceptable to the respondent). (Walton et al., 2008, p. 339)

When a problem is to be reduced to a previously proved result, the task of proving C becomes that of finding a suitable theorem A , and intervening sequence of propositions, B_1, \dots, B_n . Walton does not state critical questions for this scheme. By comparison with similar schemes, we may work out what questions are required:

- (1) What evidence is there that proposition A is true (acceptable to the respondent)?
- (2) What is the evidence showing each of the conditionals to be true?
- (3) Are there independent reasons for rejecting that the proposition C is true (acceptable to the respondent).

In Angluin's example, A is undecidable, so the first critical question cannot be answered affirmatively.

⁴The unsolvability of the halting problem is a famous result in computability theory. The quotation marks in this and all other spurious proofs are original. They may be understood as ascribing direct speech to the proponent.

Not all backward reasoning in mathematics is deductive:

Proof by importance: A large body of useful consequences all follow from the proposition in question. (Angluin, 1983, p. 17)

This piece of reasoning is interestingly ambiguous. On the one hand, it might be perceived as consequentialist practical reasoning:

Argumentation Scheme 2.2 (Argument from Positive Consequences).

Premise: If A is brought about, then good consequences will plausibly occur.

Conclusion: Therefore, A should be brought about.

Critical Questions:

- (1) How strong is the likelihood that the cited consequences will (may, must) occur?
- (2) What evidence supports the claim that the cited consequences will (may, must) occur, and is it sufficient to support the strength of the claim adequately?
- (3) Are there other opposite consequences (bad as opposed to good, for example) that should be taken into account? (Walton et al., 2008, pp. 332 f.)

Such reasoning is used to defend contested axioms or principles, such as the axiom of choice or the law of excluded middle, because as axioms they defy proof. ‘Proof by importance’ works by suggesting that an unproven conjecture must be true, because of its extensive coherence with other results. This is spurious as proof, but is the sort of hunch that, for example, was offered in support of the Taniyama-Shimura hypothesis prior to its confirmation by Andrew Wiles.⁵

Reasoning backwards from the consequences of a conjecture may also be understood as abductive reasoning. George Pólya discusses the importance of abduction for Euler in establishing his novel proof that $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$ through representation of $\sin x$ as an infinite product: ‘Euler does not reexamine the grounds for his conjecture . . . he examines only its consequences. He regards the verification of any such consequence as an argument in favor of his conjecture . . . In scientific research as in ordinary life, we believe,

⁵The hypothesis relates modular forms to elliptical curves, two seemingly remote categories of mathematical object. Its proof by Wiles was the decisive step in the proof of Fermat’s Last Theorem (see Singh, 1997, for a popular account).

or ought to believe, a conjecture more or less according as its observable consequences agree with the facts. In short, Euler seems to think the same way as reasonable people, scientists or non-scientists, usually think' (Pólya, 1954, vol. 1, p. 22). Of course, despite the importance of this informal procedure for Euler, he never mistook it for proof. Indeed he kept worrying away at the problem until he found an independent proof in more conventional terms.

Euler's inference, and some cases of Angluin's 'proof by importance', might thus be characterized by the positive form of the following abductive argumentation scheme:

Argumentation Scheme 2.3 (Argument from Evidence to a Hypothesis).

Premise: If A (a hypothesis) is true, then B (a proposition reporting an event) will be observed to be true.

Premise: B has been observed to be true [false], in a given instance.

Conclusion: Therefore, A is true [false].

Critical Questions:

- (1) Is it the case that if A is true, then B is true?
- (2) Has B been observed to be true (false)?
- (3) Could there be some reason why B is true, other than its being because of A being true? (Walton et al., 2008, pp. 331 f.)

The falsification variant is *modus tollens*, and thus uncontroversial. From a strict falsificationist perspective, the best that can be said of any hypothesis is that it has not (yet) been falsified. The defeasible acceptance of a hypothesis on the grounds of non-falsification is another application of the positive variant of this scheme. For empirical science, this may come close enough to confirmation for practical cases. In mathematics the effect is less convincing:

Proof by accumulated evidence: Long and diligent search has not revealed a counterexample. (Angluin, 1983, p. 17)

Of course this is no sort of 'proof'. If the scope of the search is explicitly stated, a more limited result could be confirmed (see §6). But counterexamples revealed only by *extraordinarily* long and diligent searches

are common enough to undercut any hope that such reasoning may hold in general. Because of such counterexamples, which have few empirical analogues, the third critical question can only be fully answered in mathematical cases by proving the hypothesis.

3. CITATION

Source-based arguments are a rich source of mathematical anxiety and humour. Yet comparatively little attention has been paid to the role of citation in mathematical practice (an important exception is Geist, Löwe & Van Kerkhove, 2010). The most numerous and diverse cases on Angluin’s list are subvarieties of deviant citation. Firstly, citations can reference things which do not exist, or which do not say what they are supposed to:

Proof by wishful citation: The author cites the negation, converse, or generalization of a theorem from the literature to support his claim.

Proof by ghost reference: Nothing even remotely resembling the cited theorem appears in the reference given. (Angluin, 1983, pp. 16 f.)

More subtly, the citation may not adequately reflect the source’s level of confidence. Angluin’s juxtaposition of the next two examples illustrates an anxiety about the credibility arbitrage behind such references:

Proof by eminent authority: ‘I saw Karp in the elevator and he said it was probably NP-complete.’

Proof by personal communication: ‘Eight-dimensional colored cycle stripping is NP-complete [Karp, personal communication].’ (Angluin, 1983, p. 16)

Other citations send the diligent on a wild goose chase which is practically (or even logically) impossible to complete:

Proof by reference to inaccessible literature: The author cites a simple corollary of a theorem to be found in a privately circulated memoir of the Slovenian Philological Society, 1883.

Proof by mutual reference: In reference A, Theorem 5 is said to follow from Theorem 3 in reference B, which is shown to follow from Corollary 6.2 in reference C, which is an easy consequence of Theorem 5 in reference A. (Angluin, 1983, p. 17)

Most transparently, the citation may refer to future activity which may never happen:

Proof by forward reference: Reference is usually to a forthcoming paper by the author, which is often not as forthcoming as at first. (Angluin, 1983, p. 17)

Proof by deferral: ‘We’ll prove this later in the course.’ (Renteln & Dundes, 2005, p. 28)

All of these different cases may be understood as ineffective applications of the same argumentation scheme, that for Appeal to Expert Opinion.

Argumentation Scheme 3.1 (Appeal to Expert Opinion).

Major Premise: Source E is an expert in subject domain S containing proposition A .

Minor Premise: E asserts that proposition A is true (false).

Conclusion: A is true (false).

Critical Questions:

- (1) Expertise Question: How credible is E as an expert source?
- (2) Field Question: Is E an expert in the field that A is in?
- (3) Opinion Question: What did E assert that implies A ?
- (4) Trustworthiness Question: Is E personally reliable as a source?
- (5) Consistency Question: Is A consistent with what other experts assert?
- (6) Backup Evidence Question: Is E 's assertion based on evidence? (Walton et al., 2008, p. 310)

Citation is indispensable in serious mathematical practice, but must be conducted with care. Walton's list of critical questions, although designed primarily for a forensic context, provides a good starting point for an analysis of what comprises such care. Good citation requires a positive answer to all six questions. All of the above examples block a satisfactory answer to one or more of these questions. For most of them, the Opinion Question could not be fully answered. (Mutual reference is an exception, but it fails the Backup Evidence Question, as do several of the others.)

All of these cases exhibit defective argumentation by the prover, whose use of the scheme is exposed as unwarranted by his inability to answer the critical questions. But the prover must also have a respondent, the audience for his proof. Pseudo-proofs can involve the argumentational shortcomings of the respondent too. For example, the respondent may be too ready to defer to the proponent's authority:

Proof by vehement assertion: It is useful to have some kind of authority relation to the audience.

(Angluin, 1983, p. 17)

Proof by adverb: ‘As is quite clear, the elementary aforementioned statement is obviously valid.’

(Cherkaev & Cherkaev, nd)

The latter case exemplifies the former; in both the prover leads his audience into employing Appeal to Expert Opinion, with him as the expert. In effect, he is appealing to his own authority. If the respondent lets him get away with it, she is neglecting her duty to challenge the assertions of the prover by asking him the appropriate critical questions. Something similar is happening in the next case:

Proof by vigorous handwaving: Works well in a classroom or seminar setting. (Angluin, 1983, p. 16)

In each of these cases the audience is browbeaten into acquiescence, something they share with another familiar scenario:

Proof by intimidation: ‘Trivial’ (Angluin, 1983, p. 16)

In cases where the result is not trivial, and perhaps not correct, this may be understood as inducing the respondent to employ a different scheme, for Argument from Danger.

Argumentation Scheme 3.2 (Argument from Danger).

Premise: If you (the respondent) bring about A , then B will occur.

Premise: B is a danger to you.

Conclusion: Therefore (on balance) you should not bring about A . (Walton et al., 2008, p. 334)

The argument induced by the dogmatic declaration ‘Trivial!’ may be unpacked as an instance of this scheme: ‘If you ask for clarification, you will appear to be a fool. Appearing a fool is a danger to you. Therefore (on balance) you should not ask.’ (Those in no danger of appearing a fool can ask devastatingly simple questions which might, from a humbler source, provoke mere condescension.) However, where the result seems trivial to the audience too, there can be no intimidation in drawing attention to its triviality; rather, this may be understood as based on a different source: (shared) intuition.

4. INTUITION

Like citation, intuition is ubiquitous, but problematic. As Reuben Hersh observes, in ‘mathematical practice, the intuitive is everywhere. . . . [It] carries a heavy load of mystery and ambiguity. Sometimes it’s a dangerous illegitimate substitute for rigorous proof. Sometimes it’s a flash of insight that tells the happy few what others learn with great effort’ (Hersh, 1997, p. 61). This ambiguity manifests in jokes:

Proof by appeal to intuition: Cloud-shaped drawings frequently help here. (Angluin, 1983, p. 17)

Proof by seduction: ‘Convince yourself that this is true!’ (Renteln & Dundes, 2005, p. 28)

The second example may be found, almost verbatim, even in so august an author as G. H. Hardy: ‘The reader will convince himself of the truth of the following assertion’ (Hardy, 1967, p. 4).

Intuition can be private, but if it is intended to convince others it must be shared; that is, it must be obvious. Hardy (crediting J. E. Littlewood) offers a partial resolution of the underlying ambiguity of intuitive obviousness: ‘One may mean “it is difficult to doubt the truth of the theorem,” . . . But sometimes we mean . . . “a moment’s reflection should not only convince the reader of the truth of what is stated, but should also suggest to him the general lines of a rigorous proof”’ (Hardy, 1967, p. 130).⁶ Thus intuitive obviousness is a knowledge claim resting on an equivocal basis. Both of Hardy’s senses may be characterized by the same scheme:

Argumentation Scheme 4.1 (Argument from Position to Know).

Major Premise: Source a is in position to know about things in a certain subject domain S containing proposition A .

Minor Premise: a asserts that A is true (false).

Conclusion: A is true (false).

Critical Questions:

- (1) Is a in position to know whether A is true (false)?
- (2) Is a an honest (trustworthy, reliable) source?

⁶Hersh also analyses the ambiguity of intuition, identifying six different senses (Hersh, 1997, pp. 61 f.). Although Hersh ranges wider, Hardy is more incisive: he subdivides the third of Hersh’s senses.

(3) Did a assert that A is true (false)? (Walton et al., 2008, p. 309)

In the second sense, the prover and (so he assumes) the respondent are in a position not only to know but to prove the ‘obvious’ result, and the critical questions may be easily answered in the affirmative. In the first sense, an adequate basis for the prover’s knowledge claim is more fugitive. It may be, as Hersh suggests, ‘reliable mathematical belief without the slightest dream of being formalized’ (Hersh, 1997, p. 61). Since he views this faculty as reliable, Hersh would still answer the critical questions affirmatively. Sceptics about such reliability would demur from one or both of the first two questions.

5. META-ARGUMENT

A prominent feature of mathematics is the reification of arguments as objects of mathematical inquiry. However, arguments which bear upon the status of other arguments comprise a variety of source-based argumentation which occurs in non-mathematical contexts. Angluin’s list contains two cases which exploit this practice in different ways.

Proof by metaproof: A method is given to construct the desired proof. The correctness of the method is proved by any of these techniques. (Angluin, 1983, p. 17)

Here the source from which the argument at issue arises is itself an argument, the presumed virtues of which would indeed transfer to the object argument, but of course, only if the method is secure.

Proof by funding: How could three different government agencies be wrong? (Angluin, 1983, p. 16)

The existence of funding, like the eminence of the author’s institution, and other ‘esteem indicators’, might be taken as indirect evidence for the soundness of his results. While this evidence is pretty weak, similar evidence can be persuasive. One empirical study presented samples of undergraduates and research mathematicians with a heuristic argument for the presence of one million consecutive sevens in the decimal expansion of π . Correctly attributing its authorship to Field’s Medallist Tim Gowers significantly increased the positive appraisal of the argument in both groups (Inglis & Mejía-Ramos, 2009). As one research subject commented, ‘We are told the argument is made by a reputable mathematician, so we implicitly assume that he would tell us if he knew of any evidence or convincing arguments to the contrary’ (Inglis & Mejía-Ramos, 2009,

p. 42). This demonstrates the perceived relevance of the author's ethos to the acceptance of his arguments, an application of the following scheme:

Argumentation Scheme 5.1 (Ethotic Argument).

Premise: If x is a person of good (bad) moral character, then what x says should be accepted as more plausible (rejected as less plausible).

Premise: a is a person of good (bad) moral character.

Conclusion: Therefore, what a says should be accepted as more plausible (rejected as less plausible).

Critical Questions:

- (1) Is a a person of good (bad) moral character?
- (2) Is character relevant in the dialogue?
- (3) Is the weight of presumption claimed strongly enough warranted by the evidence given? (Walton et al., 2008, p. 336)

Walton's version is not quite suitable for our purposes: the ethos of a must encompass epistemic instead of (or as well as) moral virtue (as Walton explores elsewhere: cf. Walton et al., 2008, p. 359). It is not Tim Gowers's blameless personal life which leads us to trust his arguments, but rather his demonstrably high standards as a working mathematician. Notoriously, academics can have deplorable ethical standards, while still producing work of genuine value. This may be because, unlike politicians, say, whose weak morals can prove more problematic, they back up their claims with independently verifiable argumentation. Does this obviate the need for consideration of their ethos, even restricted to its epistemic aspects? Only if the argument is fully worked out: if we are invited to accept aspects of the reasoning on trust, it is not unreasonable to ask whether the reasoner is trustworthy.

Ethotic Argument is closely related to Appeal to Expert Opinion. Indeed, the research subject's comments quoted above could also be construed as positive answers to some of the critical questions for the latter scheme. As the commenter is inferring the absence of counter-arguments from the absence of evidence for them, he might also be taken as employing a different form of source-based argument, an Argument from Ignorance.

6. CLOSURE

Arguments from Ignorance are typically classified as fallacious, but they can be convincing in epistemically closed domains, that is domains in which all true propositions are known.⁷ Mathematics, however, is not epistemically closed, so legitimate mathematical applications of the following scheme are scarce.

Argumentation Scheme 6.1 (Argument from Ignorance).

Major Premise: If A were true, then A would be known to be true.

Minor Premise: It is not the case that A is known to be true.

Conclusion: Therefore A is not true.

Critical Questions:

- (1) How far along has the search for evidence progressed?
- (2) Which side has the burden of proof in the dialogue as a whole? In other words, what is the ultimate *probandum* [thing to be proved] and who is supposed to prove it?
- (3) How strong does the proof need to be in order for this party to be successful in fulfilling the burden?
(Walton et al., 2008, p. 327)

This scheme is an instance of *modus tollens*, so its plausibility just turns on the acceptability of its premises. The major premise states the epistemic closure principle, which can be analyzed as a conjunction of two conditionals: (i) If A were true, then A would be known to be true by someone; (ii) If A were known to be true by someone, then A would be known to be true (by me). In much everyday social knowledge this is unproblematic: (i) is trivially true and (ii) is a straightforward claim about the speaker's position within an epistemic community.

In mathematical cases, (i) is characteristically false: there are few unknown A of which it can be said that if A were true, then A would be known to be true by someone. A notable exception is the use of an exhaustive search, whether by hand or with computer assistance. If a computer is known to have calculated the logical closure of a set of axioms, or determined which of a wide range of cases have a particular property,

⁷'Epistemic closure' is sometimes used in another sense, to specify that knowledge of a proposition implies knowledge of everything that proposition entails (or is known to entail).

then an unsuccessful search through the resultant data set is conclusive: the truth of the closure principle permits the deduction of the conclusion. Even when the closure principle is not known with certainty, it may often be practically assumed. As one science writer remarks of *Mathematica*, ‘You can enter a function that you want to integrate, and out comes the correct expression. And if it does not, you may be reasonably sure that a solution to your question simply does not exist’ (Szpiro, 2003, p. 198).

However, (ii) is often plausible in mathematics. Experts may confidently assert of many propositions in their field, that if a proof (or counterexample) were known, it would be known by them. Even non-experts can be confident about some propositions. A counter-example to Goldbach’s conjecture would be headline news. So I may safely infer from my ignorance that no such number has been found, but not of course that it does not exist. Hence the *modus tollens* stops at (ii), and the safe conclusion is not that A is not true, but that nobody knows whether A is true.

Angluin offers an interesting variant of this argument:

Proof by cosmology: The negation of the proposition is unimaginable or meaningless. Popular for proofs of the existence of God. (Angluin, 1983, p. 17)

This ‘proof’ employs a weaker closure principle: ‘If $\neg A$ were true, then $\neg A$ would be imaginable.’ The absence of an objective measure of imaginability makes this principle implausible.

Arguments from Ignorance can also be induced in the respondent:

Proof by cumbersome notation: Best done with access to at least four alphabets and special symbols.

Proof by exhaustion: An issue or two of a journal devoted to your proof is useful.

Proof by obfuscation: A long plotless sequence of true and/or meaningless syntactically related statements. (Angluin, 1983, p. 16)

These three cases encourage the respondent to pass the buck, rather than tackle a purported proof of conspicuous repugnance. The Argument from Ignorance would run ‘If there was something wrong with this proof, someone else would have noticed. Nobody seems to have noticed anything. So, it must be OK.’

7. GENERALIZATION

Much mathematical practice concerns the correct application of established rules. A notable example is the move from particular to general results. However, generalization lends itself to abuse when applied carelessly:

Proof by example: The author gives only the case $n = 2$ and suggests that it contains most of the ideas of the general proof.

Proof by omission: ‘The reader may easily supply the details.’

‘The other 253 cases are analogous.’

‘...’ (Angluin, 1983, p. 16)

These patterns of reasoning occur in both historical and contemporary mathematics. John Wallis, for instance, routinely argued from specific examples to the conclusion that a procedure may be applied indefinitely. He ‘called [this] “modus inductionis”’: later it was termed “incomplete induction.”’ (Pedersen, 1980, p. 38).

Standards of rigour have changed profoundly since Wallis’s day. But instances of proofs by example or omission still arise in ostensibly rigorous contexts, although often not without reproach. Take the dispute over Wu-Yi Hsiang’s alleged proof of Kepler’s conjecture, that the maximum density of a packing of congruent spheres in three dimensions is $\pi/\sqrt{18}$. In his review, Gábor Fejes Tóth complains that ‘we are given arguments such as “the most critical case is...” followed by a statement that “the same method will imply the general case.”’ The problem with arguments of this kind is not only that they require the reader to redo some pages of calculations, but, notoriously, that they occur at places where we expect difficulties and most frequently it is impossible to see how the same method works in the general case’ (Fejes Tóth, 1995). Both cases exemplify the following scheme.

Argumentation Scheme 7.1 (Argument from Example).

Premise: In this particular case, the individual a has property F and also property G .

Conclusion: Therefore, generally, if x has property F , then it also has property G .

Critical Questions:

- (1) Is the proposition claimed in the premise in fact true?

- (2) Does the example cited support the generalization it is supposed to be an instance of?
- (3) Is the example typical of the kinds of cases the generalization covers?
- (4) How strong is the generalization?
- (5) Do special circumstances of the example impair its generalizability? (Walton et al., 2008, p. 314)

This scheme underpins incomplete or enumerative induction. This is, of course, invalid, but still has a place in informal mathematical reasoning, for example in hypothesis formation: a is F and also G ; I can't think of anything which is F and not G ; Are all F s G ? The concern with enumerative induction is that it may be mistaken for proof. Some empirical research is worrisome: one study concluded that over 50% of trainee elementary mathematics teachers considered arguments based on specific instances to be mathematical proofs (Martin & Harel, 1989; but see Weber, 2010, for more reassuring results).

But enumerative induction is not the only application of this scheme. While Wallis might be described as employing induction, Hsiang is attempting something else, but is still employing Argument from Example. His generalizing moves are contentious because his examples are said to provide insufficient support, and to be poorly chosen, that is to fail the second and third critical questions, not because generalization is in principle non-rigorous. Where the examples are genuinely typical, and the method transparent, application of this scheme is consistent with mathematical rigour.

A special case of generalization arises with diagrammatic proof. Angluin's list contains one example that articulates the conservative attitude:

Proof by picture: A more convincing form of proof by example. Combines well with proof by omission. (Angluin, 1983, p. 17)

Defenders of the legitimacy of diagrammatic proof acknowledge the centrality of the 'Generalization Problem': 'After we have proved a theorem using a diagram, how can we legitimately generalize the configuration of this diagram to a wide (usually infinite) class of configurations, and to what class exactly?' (Kulpa, 2009, p. 81). However, they point out that similar issues arise in non-diagrammatic inference, and may be resolved in a similar fashion in the diagrammatic case, by the painstaking assurance that the particular case really does generalize, that is by answering the critical questions of the scheme for Argument from Example.

8. DEFINITION

A less contentious application of a rule is the employment of a definition. Well-chosen definitions can be pivotal to mathematical progress. At a superficial level, this choice can appear arbitrary: mathematicians have more flexibility in the concepts they take as primitive than scientists in any other domain. On the other hand, appeals to the naturalness or otherwise of definitions are frequent, and a change of definitions can revolutionize the understanding of a hitherto intractable problem. Of course, this strategy can be abused.

Proof by semantic shift: Some of the standard but inconvenient definitions are changed for the statement of the result. (Angluin, 1983, p. 17)

Instead of defining the same problem with different terms, a different (typically much easier) problem is described by the same terms. Such sharp practice is a straightforward misapplication of the scheme for Argument from Definition.

Argumentation Scheme 8.1 (Argument from Definition to Verbal Classification).

Individual Premise: a fits definition D .

Classification Premise: For all x , if x has property D , then x can be classified as having property G .

Conclusion: a has property G .

Critical Questions:

- (1) What evidence is there that D is an adequate definition, in light of other possible alternative definitions that might exclude a 's having G ?
- (2) Is the verbal classification in the classification premise based merely on a stipulative or biased definition that is subject to doubt? (Walton et al., 2008, p. 319)

If the non-standard definition D is poorly motivated, the example will fail to answer the first question; if the non-standard definition is smuggled past a classification premise only known to hold for the standard definition, the example would (also) fail to answer the second question.

9. CONCLUSION & FUTURE WORK

We have seen how a piece of mathematical folk humour draws attention to aspects of informal mathematical practice where the argumentation carries the greatest risk. We have surveyed the borderline between admissible and inadmissible instances of such risky argumentation by means of a diverse selection of argumentation schemes originally developed for non-mathematical applications (summarized in Table 1). This is further corroboration that ‘mathematical reasoning is already in accord with principles and techniques from informal logic—even if this is unnoticed by the practitioners’ (Dove, 2009, p. 150).

----- Insert Table 1 about here -----

The application of schemes to mathematical argumentation is still in its infancy, but the comparative maturity of the methodology in other domains offers inviting prospects for future research. In recent years, argumentation schemes have attracted much attention from researchers in artificial intelligence. Schemes have been implemented within systems of defeasible argumentation, and as strategies in dialogue games (Walton et al., 2008, pp. 370; 383). Progress has also been made with automated argument mining tools using schemes to find arguments, primarily in legal texts (Walton, 2011, p. 57). This approach should also lend itself to informal mathematical argumentation.

A more pessimistic consequence for the cognitive science of mathematics follows if we take seriously the suggestion that humour plays an integral role in informal reasoning. This paper has argued that informal mathematical reasoning is continuous with informal reasoning in other domains, and thereby that humour plays the same role in both. This suggests that inventive mathematics requires a good sense of humour.⁸ Some cognitive scientists argue that humour is ‘AI-complete’: that it cannot be simulated by an artificial intelligence less sophisticated than that required to simulate general human cognition (Hurley et al., 2011, p. 298). If both claims are correct, then creative mathematics, at least the way humans do it, is also AI-complete. Of course, this would not rule out the possibility of an AI-incomplete creative artificial mathematician that reasoned by methods fundamentally different from those used by humans, but if anything that may pose an even harder design challenge.

⁸More precisely, a sense of humour of a characteristic type, something some mathematicians claim to have observed (Paulos, 1991, p. 112).

ACKNOWLEDGEMENTS

I would like to thank Matthew Inglis and Simon Colton for their detailed comments on an earlier draft of this paper. Even earlier versions were delivered at conferences in Leicester, UK and Treasure Island, FL. I am grateful to the organizers and to the audiences at both conferences.

REFERENCES

- Aberdein, A. (2010). Observations on sick mathematics. In: B. Van Kerkhove, J. P. Van Bendegem & J. De Vuyst (Eds.), *Philosophical perspectives on mathematical practice* (pp. 269–300). London: College Publications.
- Angluin, D. (1983). How to prove it. *ACM SIGACT News*, 15 (1), 16–17.
- Barbeau, E. J. (2000). *Mathematical fallacies, flaws, and flimflam*. Washington, DC: Mathematical Association of America.
- Bradis, V. M., Minkovskii, V. L., & Kharcheva, A. K. (1999 [1959]). *Lapses in mathematical reasoning*. Translated by J. J. Schorr-Kon. Mineola, NY: Dover.
- Cherkaev, A., & Cherkaev, E. (n.d.). Mathematical humor. Available at: <http://www.math.utah.edu/~cherk/mathjokes.html>. Accessed March 21, 2012.
- Dove, I. J. (2009). Towards a theory of mathematical argument. *Foundations of Science*, 14 (1–2), 137–152.
- Fejes Tóth, G. (1995). Review of Hsiang, Wu-Yi, ‘On the sphere packing problem and the proof of Kepler’s conjecture.’ *Internat. J. Math.* 4 (1993), no. 5, 739–831. *Mathematical Reviews*, 95g (52032).
- Fine, G. A., & De Soucey, M. (2005). Joking cultures: Humor themes as social regulation in group life. *Humor*, 18 (1), 1–22.
- Geist, C., Löwe, B., & Van Kerkhove, B. (2010). Peer review and knowledge by testimony in mathematics. In: B. Löwe & T. Müller (Eds.), *PhiMSAMP* (pp. 155–178). London: College Publications.
- Hamblin, C. L. (1970). *Fallacies*. London: Methuen.
- Hardy, G. H. (1967). *A course of pure mathematics* (10th ed.). Cambridge: Cambridge University Press.
- Heath, T. L. (2006 [1908]). *The thirteen books of Euclid’s Elements*. New York, NY: Barnes & Noble.
- Hersh, R. (1997). *What is mathematics, really?* London: Jonathan Cape.

- Hurley, M. M., Dennett, D. C., & Adams, R. B. (2011). *Inside jokes: Using humor to reverse-engineer the mind*. Cambridge, MA: MIT Press.
- Inglis, M., & Mejía-Ramos, J. P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27 (1), 25–50.
- Krabbe, E. C. W. (2008). Strategic maneuvering in mathematical proofs. *Argumentation*, 22 (3), 453–468.
- Kulpa, Z. (2009). Main problems of diagrammatic reasoning. Part I: The generalization problem. *Foundations of Science*, 14 (1–2), 75–96.
- Martin, R. A. (2007). *The psychology of humor*. Burlington, MA: Elsevier Academic Press.
- Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, 20 (1), 41–51.
- Maxwell, E. A. (1959). *Fallacies in mathematics*. Cambridge: Cambridge University Press.
- Minsky, M. (1983). Jokes and their relation to the cognitive unconscious. In: R. Groner, M. Groner & W. F. Bischof (Eds.), *Methods of heuristics* (pp. 171–193). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Paulos, J. A. (1991). *Beyond numeracy*. New York, NY: Knopf.
- Pease, A., & Aberdein, A. (2011). Five theories of reasoning: Interconnections and applications to mathematics. *Logic and Logical Philosophy*, 20 (1–2), 7–57.
- Pedersen, K. M. (2000 [1980]). Techniques of the calculus, 1630–1660. In: I. Grattan-Guinness (Ed.), *From the calculus to set theory, 1630–1910* (pp. 10–48). Princeton, NJ: Princeton University Press.
- Pólya, G. (1954). *Mathematics and plausible reasoning* (2 vols). Princeton, NJ: Princeton University Press.
- Renteln, P., & Dundes, A. (2005). Foolproof: A sampling of mathematical folk humor. *Notices of the AMS*, 52 (1), 24–34.
- Ritchie, G. (2004). *The linguistic analysis of jokes*. London: Routledge.
- Singh, S. (1997). *Fermat's last theorem*. London: Fourth Estate.
- Szpiro, G. G. (2003). *Kepler's conjecture*. Hoboken, NJ: Wiley.
- Walton, D. N. (2011). Argument mining by applying argumentation schemes. *Studies in Logic*, 4 (1), 38–64.
- Walton, D. N., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12 (4), 306–336.

TABLE 1. Summary of reasoning types and schemes discussed

- Reasoning
 - Retroduction
 - * Argument from Gradualism
 - * Argument from Positive Consequences
 - * Argument from Evidence to a Hypothesis
- Source Based
 - Citation
 - * Appeal to Expert Opinion
 - * Argument from Danger
 - Intuition
 - * Argument from Position to Know
 - Meta-Argument
 - * Ethotic Argument
 - Closure
 - * Argument from Ignorance
- Rule Based
 - Generalization
 - * Argument from Example
 - Definition
 - * Argument from Definition to Verbal Classification