THE ABILITY TO REJECT INVALID LOGICAL INFERENCES PREDICTS PROOF COMPREHENSION AND MATHEMATICS PERFORMANCE

Lara Alcock*, Toby Bailey**, Matthew Inglis*, Pamela Docherty**
* Loughborough University, UK
** University of Edinburgh, UK

In this paper we report a study designed to investigate the impact of logical reasoning ability on proof comprehension. Undergraduates beginning their study of proof-based mathematics were asked to complete a conditional reasoning task that involved deciding whether a stated conclusion follows necessarily from a statement of the form "if p *then* q*"; they were then asked to read a previously unseen proof and to complete an associated comprehension test. To investigate the broader impact of their conditional reasoning skills, we also constructed a composite measure of the participants' performance in their mathematics courses. Analyses revealed that the ability to reject invalid denial-of-the-antecedent and affirmation-of-the-consequent inferences predicted both proof comprehension and course performance, but the ability to endorse valid modus tollens inferences did not. This result adds to a growing body of research indicating that success in advanced mathematics does not require a normatively correct material interpretation of conditional statements.*

Key words: Conditional Inference, Logical Reasoning, Proof Comprehension, Undergraduate Mathematics Education

Introduction

Mathematics and logical reasoning are seen as closely related. It is widely believed that study of mathematics develops general logical reasoning skills (e.g., NCTM, 2000), and that correct logical reasoning is important for study of advanced mathematics: transition-to-proof textbooks commonly deal with the topic explicitly. But is this really the case? Recent research has revealed that mathematics and logical reasoning are related, but that this relationship is not straightforward: it is not the case that experienced mathematicians uniformly conform to normatively correct interpretations of conditional statements (e.g., Inglis & Simpson, 2006). This raises questions about what we need to teach and about which failures of reasoning should worry us. In this paper we take up this discussion, arguing that mathematics education does not lead students to normatively correct reasoning, but does nevertheless develop the logical reasoning skills that students need for advanced mathematics. We begin by reviewing arguments and evidence on expert and novice reasoning with *conditional statements* of the form "if p then q."

Reasoning with Conditional Statements

Consider a conditional statement about an imaginary letter-number pair:

"If the letter is X then the number is 1."

Researchers have investigated patterns of responses to four possible inferences from this statement plus a related assertion; these inferences are listed in Table 1.

Assertion	Inference	Inference type
The letter is X (p)	The number is 1	*Modus ponens*
The letter is not X (not-p)	The number is not 1	*Denial of the antecedent*
The number is 1 (q)	The letter is X	*Affirmation of the consequent*
The number is not 1 (not-q)	The letter is not X	*Modus tollens*

Table 1: Four possible inferences from the conditional statement plus a related assertion.

In formal logic, where conditional statements are interpreted as *material conditionals*, modus ponens (MP) and modus tollens (MT) inferences are defined to be valid, and denial-of-the-antecedent (DA) and affirmation-of-the-consequent (AC) inferences are defined to be invalid. This is the interpretation taught in textbooks and in transition-to-proof-courses. It is not, however, the interpretation commonly made by people without formal training. In everyday life, it has been argued that a *defective conditional* interpretation is more common (e.g., Quine, 1966). Under this interpretation, the conditional statement is seen as stating only that the consequent follows given that the antecedent is true, meaning that the statement is irrelevant in cases in which the antecedent is not true. The picture is further complicated by the common everyday interpretation of a conditional statement "if p then q" as the biconditional statement "p if and only if q" (e.g., Epp, 2003). These three interpretations are compared in Table 2.

p	q	if p then q (biconditional)	if p then q (defective)	if p then q (material)
T	T	T	T	T
T	F	F	F	F
F	T	F	irrelevant	T
F	F	T	irrelevant	T

Table 2: Comparison of biconditional, defective conditional and material conditional interpretations (T – true; F – false).

The corresponding responses to the four inference types are given in Table 3. Under a biconditional interpretation, p and q are seen as simply "going together," so that either both are true or both are false. This corresponds to endorsement of all four inferences. Under a defective interpretation, only modus ponens is endorsed, since the other three inferences involve assertions for which the conditional statement is seen as irrelevant. The material interpretation corresponds to the normatively correct responses as listed above.

Inference type	biconditional	defective	material
Modus ponens	endorse	endorse	endorse
Denial of the antecedent	endorse	reject	reject
Affirmation of the consequent	endorse	reject	reject
Modus tollens	endorse	reject	endorse

Table 3: Comparison of biconditional, defective conditional and material conditional interpretations.

It might be natural, then, to see the material interpretation as the most sophisticated, and to believe that mathematics educators should help students develop toward this interpretation and should be concerned if it is not attained. But is it true that a material interpretation is important for mathematical success?

Evidence on Mathematical Education and Logical Reasoning

It has been argued that a material interpretation is needed for advanced mathematics; that certain forms of indirect reasoning are not accessible without it (Durand-Guerrier, 2003). However, it has also been argued that at lower educational levels a defective interpretation is more appropriate, "since in school mathematics, students have to appreciate the consequence of an implication when the antecedent is taken to be true" (Hoyles & Küchemann, 2002, p. 196). Indeed, evidence indicates that mathematical study develops conditional reasoning skill, but develops it toward a defective rather than a material interpretation. In a sample of students in the UK, where compulsory education ends at 16, those studying mathematics in the first non-compulsory year were found to change in their responses to a conditional reasoning task more than did an equivalent population studying English literature (and not mathematics). The mathematics students became more likely to reject AC inferences and DA inferences, but also more likely to reject MT inferences (Attridge & Inglis, 2013). Might this mean that their education did a disservice to those who went on to study undergraduate mathematics? Does pre-proof mathematical education teach students a better but still inadequate interpretation of the conditional, and does this cause problems when they come to study proof?

Surprisingly, evidence from work with professional mathematicians suggests that it might not, because mathematicians do not reliably make the material interpretation either. In a study of mathematicians' responses to the Wason Selection Task, Inglis and Simpson (2006) showed that professional mathematicians behave differently from members of a general educated population: they are not tempted by AC and DA inferences, but neither do they reliably consider a relevant MT inference. Perhaps, then, a defective interpretation is perfectly adequate for success even in proof-based mathematics. Our data supports this suggestion, as described below.

Methods

Participants in our study were 112 students in a first year, second semester undergraduate mathematics class on problem solving and proving (the equivalent of a transition-to-proof course). All had taken a linear algebra course in the previous semester (this included theorems and proofs but treated these quite lightly) and were concurrently enrolled on a course in calculus (this included some proofs and some calculations involving limits, but epsilon-delta techniques appeared only briefly). All were spending 50% of their total time over the year in these mathematics classes, and for almost all this was a compulsory component of a degree programme with "mathematics" in the title. In workshop sessions in week 8 of the 11-week problem solving and proofs course, participants were asked to complete a conditional reasoning test, and to read and answer comprehension questions on a previously unseen proof. They completed both tasks individually and in silence.

The conditional reasoning test (adapted from Evans, Clibbens & Rood, 1995) comprised 16 items of the form shown in Figure 1. There were four items for each type of inference (MP, AC, DA and MT), and instructions asked participants to decide whether the conclusion necessarily follows and to indicate their answer by placing a check mark in the appropriate circle. Participants were given ten minutes to complete the task, and the order of the items was randomised for each participant. For analysis purposes, a count out of four was constructed for each inference type, where each point indicated an instance in which the participant agreed with an inference of the relevant type.

If the letter is J then the number is not 2.
The number is 7.
Conclusion: The letter is J.
○ YES.
○ NO.

Figure 1: A conditional reasoning test item (an AC item).

The proof comprehension task involved a proof that the product of two primes is not abundant (i.e., that the product is not less than the sum of its proper factors). Participants were asked to study the proof carefully and then to answer a proof comprehension test based on the model developed by Mejía-Ramos, Fuller, Weber, Rhoads and Samkoff (2012). This test comprised ten multiple-choice items, each of which had two distractors and one correct answer (the proof and test are too long to reproduce here, but full copies will be provided at the talk if this paper is accepted). Participants were allowed 15 minutes for this task.

Finally, we obtained the participants' examination scores in their calculus, linear algebra and problem solving and proofs courses (all three courses had some coursework together with a final individual summative examination worth 85% of the course grade). The average of these scores was used as a measure for performance in core mathematics courses.

Results

Table 4 presents the descriptive statistics for all six measures, showing the minimum and maximum number of inferences of each type endorsed, as well as the associated means and standard deviations.

Measure (theoretical max)	Min	Max	Mean	Std. Dev.
MP inferences endorsed (4)	2	4	3.68	0.541
DA inferences endorsed (4)	0	4	1.06	1.085
AC inferences endorsed (4)	0	4	1.42	1.271
MT inferences endorsed (4)	0	4	2.68	1.050
Proof Comprehension (10)	3	10	7.29	0.182
Math Course (100)	22	95	61.86	16.426

Table 4: Descriptive statistics.

We note that the MP counts were close to ceiling. This is to be expected, but it renders this measure inappropriate for use in regression models, so we omit it in the following analyses. The counts for the remaining conditional reasoning measures show considerable variability – participants on average endorsed more than one of each of the invalid DA and AC inferences, and rejected more than one valid MT inference. The proof comprehension scores were generally high, and the range and average of the mathematics performance scores were typical in the national context in which the study took place.

Table 5 presents regression models with the DA, AC and MT counts as independent variables and with (a) proof comprehension score and (b) mathematics performance score as the dependent variables. Figure 3 shows the means of the proof comprehension and mathematics performance scores for participants with different DA, AC and MT counts, together with lines of best fit for cases in which the count is a significant predictor.

(a)

R^2	Predictors	β	p
.171***	DA	-.234	.029
	AC	-.234	.031
	MT	.003	.969

(b)

R^2	Predictors	β	p
.261***	DA	-.209	.041
	AC	-.342	.001
	MT	.084	.321

Table 5: Regression models predicting (a) proof comprehension score and (b) mathematics performance; ***p<.001.

.

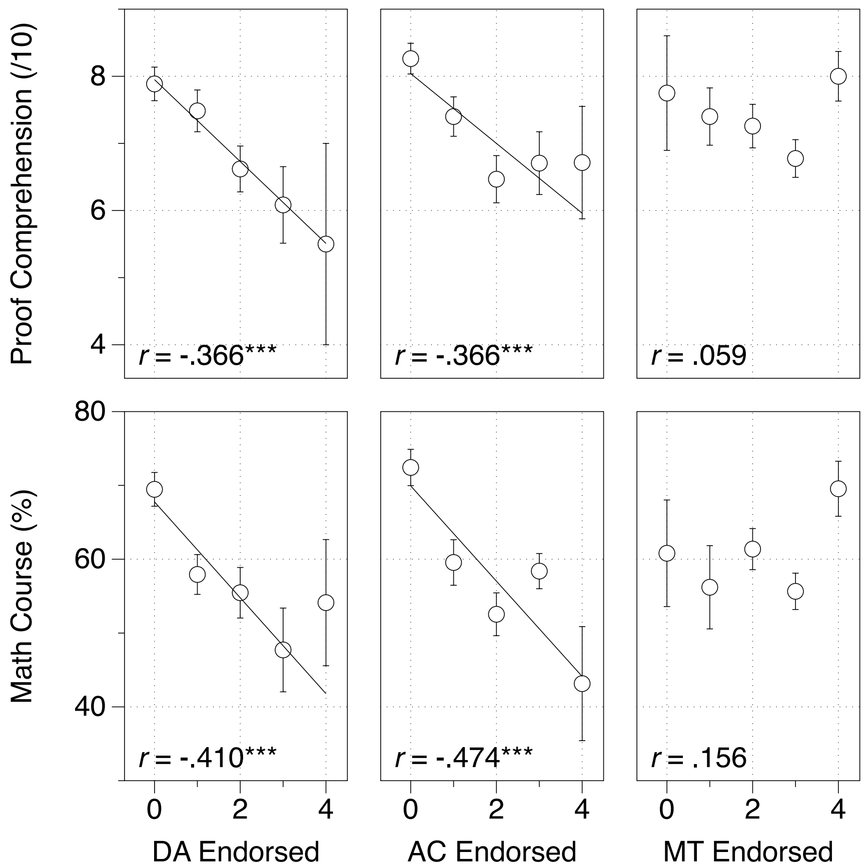

Figure 2: Proof Comprehension test and Math Course score means and correlation coefficients for participants endorsing different numbers of DA, AC and MT inferences; error bars show ±1 SE of the mean, *** $p < .001$.

In both models, the DA and AC counts were significant predictors with negative coefficients: participants who rejected more DA and more AC inferences performed better both on the proof comprehension test and in their mathematics courses. In both models, the MT score was not a significant predictor: endorsement of MT inferences, which are valid under the normatively correct material interpretation of the conditional, did not have a systematic effect on either outcome measure.

Discussion

The material interpretation of the conditional is normatively correct and is taught in standard undergraduate mathematics. However, this study adds to a growing set of results suggesting that full conformity with its entailments is not necessary for mathematical success. While the ability to reject invalid DA and AC inferences does appear to predict success in proof comprehension and in undergraduate-level courses, the ability to reliably endorse valid MT inferences does not.

One obvious limitation of this study is that it involved a comprehension test for only a single proof. Concern about this should be mitigated by the second regression model in which performance across three core mathematics modules showed a similar pattern; if a lack of ability to endorse MT inferences were a serious problem, we would expect it to appear as a significant predictor in this model. However, these results leave open the possibility that a defective interpretation of the conditional is a disadvantage under some specific circumstances. Perhaps, for instance, students with this interpretation are less able to understand contradiction or contraposition arguments. All such arguments relative to a statement of the form "if *p* then *q*" involve a step at which one establishes not-*q* and uses this to conclude not-*p*. Thus, we might expect them to be less well understood by students who do not readily endorse MT inferences. This could be investigated, although we suggest that such work should be done in parallel with further investigation of how expert mathematicians process such arguments. Recall that mathematicians do not reliably consider relevant MT inferences under all circumstances, so there might not be a straightforward link between reasoning about single abstract conditional statements and understanding this structure as it is used in proofs. Indeed, Inglis and Simpson (2009) suggest that the equivalent of an MT inference can be constructed given a defective interpretation of the conditional statement "if *p* then *q*" and the assertion not-*q*: they might suppose *p*, conclude *q* by MP, note that this contradicts the assertion, and conclude that their supposition of *p* was incorrect. This is a somewhat long chain of reasoning, but that very fact might account for all of the results: if this is the mechanism typically used, we would expect that neither mathematicians nor students would endorse all straightforward MT inferences by simple recognition, but that experienced mathematicians and more successful students would be better able to reach correct conclusions by correctly reasoning through the whole chain.

Prior to such investigations, we do not suggest that we should stop teaching mathematics students the material interpretation of the conditional. However, we do suggest that we should not be too concerned if undergraduate students do not develop to a point at which they reliably endorse MT inferences, because it appears that they may not need to.

References

Attridge, N., & Inglis, M. (2013). Advanced mathematical study and the development of conditional reasoning skills. *PLoS ONE*, *8*, e69399.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics*, *53*, 5-34.

Epp, S. (2003). The role of logic in teaching proof. *American Mathematical Monthly*, *110*, 886-899.

Evans, J. St. B. T., Clibbens, J., & Rood, B. (1995). Bias in conditional inference: Implications for mental models and mental logic. *Quarterly Journal of Experimental Psychology*, *48A*, 644-670.

Hoyles, C., & Küchemann, D. (2002). Students' understanding of logical implication. *Educational Studies in Mathematics*, *51*, 193-223.

Inglis, M., & Simpson, A. (2006). The role of mathematical context in evaluating conditional statements. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková, (eds.), *Proceedings of the 30th International Conference on the Psychology of Mathematics Education* (Vol. 3, pp. 337-344). Prague, Czech Republic. IGPME.

Inglis, M., & Simpson, A. (2009). The defective and material conditionals in mathematics: Does it matter? In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis, (eds.), *Proceedings of the 33rd International Conference on the Psychology of Mathematics Education* (Vol. 3, pp.225-232), Thessaloniki, Greece: IGPME.

Mejía -Ramos, J.-P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*, 3-18.

NCTM (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Quine, W. V. O. (1966). *Methods of Logic (second edition)*. Routledge & Kegan Paul: London.