Logic in Introduction-to-Proof Students: Mathematical Truth Table Tasks

Lara Alcock
Loughborough University, UK

Mathematics uses standard logic, in which the truth value of the conditional 'if A then B' is completely determined by the truth values of A and B. Specifically, when A is false, 'if A then B' is, by definition, true. This makes the mathematical conditional out of line with natural language, where research using truth table tasks shows that conditionals are often considered irrelevant in false-antecedent cases. What does this mean for mathematics undergraduates? Do they interpret conditionals according to standard logic, or do their responses align with everyday interpretations? Do they make consistent interpretations, or do these vary by content? This report contributes a first investigation of this issue.

Keywords: conditionals, if-then, logic, reasoning, truth tables

Introduction

In mathematics, conditionals – *if-then* statements – are ubiquitous. Many theorems and conjectures are naturally expressed as *universal* or *generalized* conditionals of the form '$\forall x$, if $A(x)$ then $B(x)$' (Durand-Guerrier, 2003). This is so well understood that the quantification is often left implicit (Shipman, 2016), as in these example theorems:

- If $f: \mathbb{R} \to \mathbb{R}$ is differentiable at a, then f is continuous at a;
- If G is a finite group and H is a subgroup of G, then $|G| = |H||G: H|$;
- If p is a prime number and $p \nmid a$, then $a^{p-1} \equiv 1(\text{mod } p)$.

The interpretation of universal conditionals is underpinned by the standard mathematical interpretation of *if-then* as the *material conditional*, a truth-functional logical connective (Edgington, 2020). For propositions A and B, the truth value of 'if A, then B' is defined for every possible combination of truth values of A and B, and is completely determined by those truth values per the standard truth table – see Table 1.

Table 1. Defining truth table for if-then as a truth functional logical connective.

A	B	if A then B
T	T	T
T	F	F
F	T	T
F	F	T

Truth functionality is necessary for *if-then* to interact with the other connectives *and*, *or* and *not* in a way that enables unambiguous evaluation of logical equivalence and argument validity (Smith, 2020). And the standard truth table is the only viable option (Edgington, 2020). We want a universal conditional 'if $A(x)$ then $B(x)$' to be false only when there exist cases corresponding to the TF row, with $A(x)$ true but $B(x)$ false. We want it to be true if there are no such cases, compelling us to define it as true for TT, FT and FF rows. This interpretation confirms validity for standard argument forms. For instance, *modus tollens* inferences from *if A then B* and *not-B* to *not-A* are valid because the premises are true together only in the FF row of the table, where A is false (Velleman, 2006). However, the truth-functional conditional is psychologically unnatural. It defines a conditional as true whenever its antecedent is false, where natural-language conditionals are often considered irrelevant in such cases (Nickerson, 2015). This issue is the focus of this report, which describes the development and use of mathematical *truth table tasks*.

Theoretical Background

Conditionals and their interpretation have been studied in mathematics education (Dawkins & Norton, 2022; Inglis & Attridge, 2016) and more generally in psychology (Nickerson, 2015), philosophy (Edgington, 2020) and linguistics (Douven et al., 2020). Reasoning *from* conditionals is often investigated using *conditional inference tasks*, which have been used to study mathematics students' reasoning about both abstract content (Attridge & Inglis, 2013) and everyday and mathematical content (Alcock et al., under review; Alcock & Inglis, under review). Interpretation *of* conditionals is often studied using *truth table tasks*, which have not so far been used with mathematics students or with mathematical content.

Truth table tasks come in various forms. One simple type presents a conditional and asks participants to generate examples that are true and examples that are false. People commonly generate TT cases for 'true' and TF cases for 'false'; a minority of participants additionally generate false-antecedent cases for 'true' (Nickerson, 2015).

A more standard type presents participants with a conditional and with cases corresponding to each truth table row; the task is to judge the truth value of the conditional in each case. Tasks vary in their response options: participants might be forced to choose between 'true' and 'false' or might be allowed a third option of 'irrelevant', 'indeterminate' or similar (Nickerson, 2015). When offered this third option, people typically use it for one or both false-antecedent cases. For instance, Newstead et al. (1997) asked participants to state whether each presented case 'supports', 'contradicts' or 'tells us nothing about' the relevant conditional. Across three studies, they found that people almost always judged that TT cases support the conditional and TF cases contradict it. For FT cases, responses were split almost equally between contradicts and tells us nothing; for FF cases, responses were split almost equally between supports and tells us nothing. This effectively gives the 'truth table' in Table 2.

Table 2. Common responses in truth table tasks, Newstead et al. (1997).

A	B	if A then B
T	T	T
T	F	F
F	T	F or I
F	F	T or I

Tasks also vary in semantic content. Newstead et al. systematically varied their content using conditional promises, threats, tips, warnings, and conditionals that they called temporal, causal, and universal. Universals (e.g., 'If the student is doing economics, then he is a socialist') are arguably closest to mathematical conditionals because they are effectively about set inclusion, though of course such claims would be treated as debatable – the bivalent logic of mathematics does not usually apply to everyday content (Oaksford & Chater, 2020). For universal conditionals, Newstead et al. reported higher proportions of tells-us-nothing responses, at 64% for FT and 63% for FF cases respectively; this meant that 46% of responses followed a TFII pattern sometimes referred to as a *defective implication* interpretation (following Wason &

Johnson-Laird, 1969). The full set of patterns considered by Newstead et al. were termed *implication* (TFTT, as in the truth-functional conditional), *equivalence* (TFFT, treating a conditional as a biconditional), *defective implication* (TFII), and two patterns based on their data, *defective biconditional* (TFFI) and *Pattern X* (TFIT). The implication pattern was extremely rare: interpretations of everyday conditionals do not align with standard logic, as is well documented across tasks of various types (e.g., Oaksford & Chater, 2020).

What does this mean for students learning mathematics? When faced with mathematical content, do they import everyday interpretations, or does their reasoning align with standard logic? This study was designed to find out.

Method

Designing a truth table task with 'realistic' mathematical content is nontrivial. Mathematical conditionals are easy to come by – as noted in the Introduction, many theorems can be expressed in this form. But theorems, of course, are true, whereas a truth table task requires cases for all four truth-value combinations, TT, TF, FT and FF. This means that both the conditional and its converse must permit counterexamples. I therefore used simple mathematical content to construct conditionals that sound sensible because a) the semantic contents of their antecedents and consequents involved related concepts, as is pragmatically expected in everyday conditionals (Grice, 1989), and b) the first case to come to mind would likely be a TT case. To keep the tasks short while providing at least some mathematical variety, I used three conditionals:

- If n is a positive odd number, then n is prime.
- If rectangle R has area 12 square units, then its perimeter is 14 units.
- If x is a whole number, then $x^2 > x$.

I used all three conditionals in each of three truth table tasks (using the word 'claim' rather than 'conditional' in these tasks to make the language more accessible).

- In the *generate* task, participants were asked to give one or more examples that would make each conditional true and one or more than would make it false;
- In the *true-false* task, participants were shown the conditional with TT, TF, FT and FF cases, and asked to judge the conditional true or false for each one;
- In the *true-false-irrelevant* task, participants were shown the conditional with TT, TF, FT and FF cases, and asked to judge the conditional true, false or irrelevant for each one.

For the true-false and true-false-irrelevant tasks, the provided cases are shown in Table 3. Rectangles were shown diagrammatically – see the example true-false-irrelevant task in Figure 1.

Table 3. TT, TF, FT and FF cases provided for each conditional in the true-false and true-false-irrelevant tasks; for task presentation format see the appendix.

Conditional	TT	TF	FT	FF
If n is a positive odd number, then n is prime.	3	9	2	8
If rectangle R has area 12 square units, then its perimeter is 14 units.	3×4	2×6	1×6	2×3
If x is a whole number, then $x^2 > x$.	3	1	3/2	1/2

Consider this claim:

If n is a positive odd number, then n is prime.

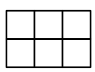
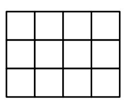
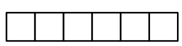
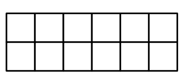
For each n-value listed below, would you say the claim is true, false, or irrelevant?

1. $n = 3$ A
2. $n = 8$ D
3. $n = 2$ C
4. $n = 9$ B

Consider this claim:

If rectangle R has area 12 square units, then its perimeter is 14 units.

For each rectangle shown below, would you say the claim is true, false, or irrelevant?

1.  D
2.  A
3.  C
4.  B

Consider this claim:

If x is a whole number, then $x^2 > x$.

For each x-value listed below, would you say the claim is true, false, or irrelevant?

1. $x = \frac{3}{2}$ C
2. $x = 3$ A
3. $x = 1$ B
4. $x = \frac{1}{2}$ D

Figure 1. Example true-false-irrelevant task page (small letters to the right track randomization).

For each task, the order in which the conditionals appeared was counterbalanced. For the true-false and true-false-irrelevant tasks, the order in which the cases appeared for each conditional was randomized. Participants were first-year UK mathematics undergraduates in the third week of an introduction-to-proof course (the researcher had no involvement in the course). Each participant received one of the three tasks on paper, and papers were arranged so that students in adjacent seats would receive different tasks. The task took five minutes. For the generate task, 29 participants consented to have their data used. For the true-false task, 30 consented but one clearly misunderstood the task and two did not respond for all three conditionals so, for simplicity, data from 27 is presented. For the true-false-irrelevant task, 26 participants consented but one did not respond for all conditionals, so data from 25 is presented.

Results

The headline result is that in judging truth values for mathematical conditionals, introduction-to-proof students adhere much more closely to the defective TFII interpretation than to the truth-functional TFTT interpretation of standard logic. That is, they interpret mathematical conditionals as true in TT cases, false in TF cases, and – where permitted and indeed sometimes where not permitted – as irrelevant in FT and FF cases. Moreover, content appears to affect interpretations, and a substantial minority of students do not implement a consistent interpretation across contents. I report results for the generate task, then break down the results for the true-false and true-false-irrelevant tasks by conditionals then by participants.

Generate Task

Results from the generate task were clear and in line with common responses to tasks with everyday content (e.g., Nickerson, 2015): for examples that would make the conditional true, participants gave TT cases; for examples that would make it false, they gave TF cases. A small number made interpretation or calculation errors (e.g. failing to notice the whole number requirement for x and providing a range such as $0 < x < 1$). But most responded predictably: Table 4 gives examples that were most commonly listed first.

Table 4. Commonest first responses for examples that make each conditional true/false, with percentage of participants giving that response.

Conditional	True	False
If n is a positive odd number, then n is prime.	3 (66%)	9 (66%)
If rectangle R has area 12 square units, then its perimeter is 14 units.	3×4 (86%)	2×6 (69%)
If x is a whole number, then $x^2 > x$.	2 (55%)	1 (66%)

True-False Task

Results for the true-false task – see Table 5 – showed that for all three conditionals, small minorities of participants aligned with the material interpretation, with response pattern TFTT. But most went against it, either choosing ‘false’ for both false-antecedent cases (TFFF) or subverting the task by introducing ‘irrelevant’ as an extra option and applying it to those cases (TFII). No participant gave a biconditional (TFFT) response. For the conditional ‘If x is a whole number, then $x^2 > x$’, there were fewer TFFF responses but several TFTF responses, judging the conditional true for the case $x = 3/2$. This differs from the common everyday responses summarized in Table 2, but the task required written one-word responses only, so I have no qualitative data on why this might be. I hope this might generate discussion at the conference.

Table 5. Numbers giving the commonest response patterns for each conditional (26 participants).

Conditional	TFFF	TFII	TFTF	TFTT
If n is a positive odd number, then n is prime.	17	3	1	4
If rectangle R has area 12 square units, then its perimeter is 14 units.	17	4	1	1
If x is a whole number, then $x^2 > x$.	11	3	8	3

Analysing by participants reflected these results in that:

- 8 responded TFFF for all three conditionals;
- 5 responded TFFF for the n and R conditionals but TFTF for the x conditional;
- 3 responded TFII for all three conditionals;
- 1 responded TFTT for all three conditionals.

Of the remaining ten participants, nine were inconsistent – some gave different response patterns for all three conditionals. As in Dawkins and Cook's (2017) work on disjunctions, this shows that students do not necessarily recognize logical structures or apply these in the same way across different mathematical content. One participant responded IFII for all three conditionals, apparently unwilling to assign a truth value in any but the clearly false case. Although unusual in these data, this is in line with interpretations in which conditionals are not considered true unless there is an inferential link between antecedent and consequent (e.g., Douven et al., 2020).

True-False-Irrelevant Task

Results for the true-false-irrelevant task – see Table 6 – showed that for all three conditionals, participants who were offered the 'irrelevant' option commonly used it, leading to response pattern TFII. No participant offered this option responded according to the material conditional (TFTT), and again no participant gave a biconditional (TFFT) response. As in the true-false task, there were a few TFTF responses for the x conditional.

Table 6. Numbers giving the commonest response patterns for each conditional (25 participants).

Conditional	TFFF	TFII	TFTF	TFTT
If n is a positive odd number, then n is prime.	2	20	0	0
If rectangle R has area 12 square units, then its perimeter is 14 units.	4	17	0	0
If x is a whole number, then $x^2 > x$.	1	16	5	0

Analysing by participants reflected these results in that:

- 14 responded TFII for all three conditionals;
- 1 responded TFII for the n and R conditionals but TFTF for the x conditional.

Of the remaining ten participants, all were inconsistent across the conditionals – again, some gave different response patterns for all three.

Discussion

This is the first study to design and use a mathematical version of a truth table task to investigate interpretations of mathematical conditionals. Its main result is that for simple mathematical content, the interpretations made by introduction-to-proof students align much more closely with typical interpretations of everyday conditionals than with standard mathematical logic. That is, students most commonly think of conditionals as true in TT cases, false in TF cases, and irrelevant in FT and FF cases. When forced to choose between true and false for false-antecedent cases, most pick 'false', giving a TFFF rather than a TFTT response. Many students are not consistent in their interpretations across content (cf. Dawkins & Cook, 2017). Finally, content seems to have an effect, with the conditional 'If x is a whole number, then $x^2 > x$' prompting some TFTF responses.

One question is how much this matters. Although the truth-functional material conditional underlies mathematical interpretations of logical equivalence and validity, 'real' mathematics rarely involves propositional conditionals. Consequently, it is possible to provide an introduction to proof that does not address propositional logic, instead focusing on the distinction between a (universal) conditional and its converse – several commonly recommended books do this (Alcock & Sa, 2025). Moreover, a defective interpretation does not appear fatal to strong mathematical reasoning. Although a material interpretation is technically necessary for validity of modus tollens inferences (Durand-Guerrier, 2003), these also submit to a longer but valid workaround ('if A were true, then B would be true, but B is not true, so A cannot be true'), and evidence from other tasks too indicates that this interpretation is common among successful undergraduates (Inglis & Simpson, 2009).

One way to approach an informed view on this question would be to find out how professional mathematicians respond to truth table tasks. It could be that all give TFTT responses, in which case an argument about developing expertise suggests that teaching the material conditional is worth the effort. But mathematicians, too, have little cause to consider propositional conditionals in their day-to-day work, so it is possible that they have little practice in thinking about these and that some would be influenced by everyday language and respond according to the defective TFII interpretation. In that case, the question becomes more nuanced – if the truth-functional interpretation is mathematically 'correct' but not necessary for a successful mathematical career, perhaps teaching it should not be a priority. By the time of the conference, I will have evidence on mathematicians' responses to these tasks.

In the meantime, my current view is that mathematics undergraduates should encounter propositional logic as part of exposure to the foundations of their subject (Alcock, 2025). However, we should expect that they will find it unnatural, and we should think carefully about whether and how our explanations leverage everyday content. In some cases, such content might support students in reaching appropriate interpretations (Epp, 2003). In others, it might encourage the reverse, or at least it might fail to raise the right questions. For instance, many textbooks that do cover the material conditional do so by presenting an everyday conditional (e.g. 'If it stops raining by Saturday, then I will go to the football game', Lay, 2014, pp.4-5); they invite students to consider when they would consider it false, then argue that it should be considered true in all other cases. As Alcock and Sa (2025) point out, this fails to engage with the irrelevance problem. To remedy this, the psychology literature suggests that an intermediate step might be useful: although people consider conditionals irrelevant for false-antecedent cases, they are often willing to concede that conditionals are *compatible* with those cases (Nickerson, 2015). The present research with mathematical content did not test this, but a follow-up is underway that will replicate the true-false and true-false-irrelevant tasks and add one testing judgements of compatibility. Again, I will hope to share the data at the conference.

Acknowledgments

This work was funded by Leverhulme Trust Research Fellowship RF-2022-155 titled "Does Mathematics Develop Logical Reasoning?".

References

Alcock, L. (2025). Understanding mathematical conditionals: An educational perspective informed by philosophy, linguistics and psychology. *Journal of Mathematical Behavior, 78*, 101233.

Alcock, L. & Sa, R. (2025). How do introduction-to-proof textbooks explain conditionals and implications? *International Journal of Research in Undergraduate Mathematics Education*. https://doi.org/10.1007/s40753-024-00263-1

Dawkins, P. C. & Cook, J. P. (2017). Guiding reinvention of conventional tools of mathematical logic: students' reasoning about mathematical disjunctions. *Educational Studies in Mathematics*, *94*, 241–256.

Dawkins, P. C. & Norton, A. (2022). Identifying mental actions for abstracting the logic of conditional statements. *Journal of Mathematical Behavior*, *66*, 100954.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantics. *Thinking & Reasoning*, *26*, 311–351.

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.

Edgington, D. (2020). Indicative conditionals. In E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2020). https://plato.stanford.edu/archives/fall2020/entries/conditionals/.

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

Grice, H. P. (1989). *Studies in the way of words*. Boston, MA: Harvard University Press.

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

Lay, S. R. (2014). *Analysis with an introduction to proof* (5th edition). Pearson New International.

Newstead, S. E., Ellis, M.C., Evans, J. St. B. T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking & Reasoning*, *3*, 49–76.

Nickerson, R. S. (2015). *Conditional reasoning: The unruly syntactics, semantics, thematics, and pragmatics of "if"*. Oxford University Press.

Oaksford, M. & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305–330.

Shipman, B. A. (2016). Subtleties of hidden quantifiers in implication. *Teaching Mathematics and its Applications*, *35*, 41–49.

Smith, P. (2020). *An introduction to formal logic: Second Edition*. Cambridge University Press.

Velleman, D. J. (2006). *How to prove it: A structured approach*. Cambridge University Press.

Wason, P. C. & Johnson-Laird, P. N. (1969). Proving a disjunctive rule. *Quarterly Journal of Experimental Psychology*, *21*, 14–20.