

Open Encyclopedia of Cognitive Science

Deductive Reasoning

Ruth M.J. Byrne¹

¹Trinity College Dublin, University of Dublin

MIT Press

Published on: Oct 07, 2026

URL: <https://oecs.mit.edu/pub/8tkz4eis>

License: [Creative Commons Attribution-NonCommercial 4.0 International License \(CC-BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/).

An enduring question that has preoccupied humankind for centuries is whether humans are rational beings. Are people capable of abstract hypothetical thought, and can they reason well from the information available to them to infer a good conclusion? When people reason deductively, they try to reach a conclusion that must be true provided the information on which it is based is true. Experiments have been carried out on a wide variety of deductive inferences, and conditional inferences are one of the most commonly studied. People find some conditional inferences easy to make and others very difficult. Their mistakes are informative about the sorts of mental representations and cognitive processes that they rely on to make inferences.

History

Human deductive reasoning has been studied since at least Aristotle's proposals over 2,000 years ago, and the first psychological experiments on it were carried out over 100 years ago. Early experiments examined categorical syllogisms such as, "All the artists in the room are beekeepers, some of the beekeepers in the room are chemists, are some of the chemists in the room artists?" (see [Khemlani & Johnson-Laird, 2012](#) for a meta-analysis). Early experiments also examined relational inferences such as, "The book is to the left of the lamp, the plant is to the right of the book, is the plant to the right of the lamp?" Nowadays, many contemporary studies examine hypothetical inferences from conditionals, such as, "If Lisa has an essay to write, then she studies late in the library. In fact, Lisa has an essay to write, does she study late in the library?" (see [Nickerson, 2015](#) for a review). Conditional reasoning is central to many aspects of thinking, from the formation of plans and hypotheses to the consideration of causes and consequences and assessment of possibilities and probabilities [see [Causal Reasoning](#)].

Conditional inferences were first studied extensively in the selection task (see [Table 1](#)). Participants are told a rule such as, "If a card has a vowel on one side, then it has an even number on the other side," and they are shown four cards with letters and numbers, for example, A, D, 4, and 7. Their task is to select the cards they need to turn over to discover whether the rule is true or false (see [Ragni et al., 2018](#) for a review). Most participants choose the A card, but few choose the 7 card. The A card can show the rule is false if it does not have an even number on the other side, but so too can the 7 card show the rule is false if it has a vowel on the other side. Participants make different inferences when the rule is realistic, such as, "If an envelope is sealed, then it has a 50 lire stamp on it." When they are shown a sealed envelope, an unsealed one, one with a 50 lire stamp, and one with a 40 lire stamp, most choose the sealed envelope and the envelope with the 40 lire stamp. The effects of content and context on deductive reasoning have been examined in many ways since. Contemporary experiments examine how people assess the truth and validity of conditional inferences, detect inconsistencies, and fall prey to illusory inferences. They examine many aspects of the inferences people make, including the sorts of conclusions they reach, their beliefs in their conclusions, the length of time they take to make them, and the typical mistakes they make.

Table 1

The Selection Task, With an Example of Rules and Instructions for Abstract and Realistic Content

Table 1				
Abstract rule: If a card has a vowel on one side, then it has an even number on the other side.				
Instructions: Select the cards that you need to turn over to discover whether the rule is true or false.	A	D	2	7
Realistic rule: If an envelope is sealed, then it has a 50 lire stamp on it				
Instructions: Select the envelopes that you need to turn over to find out whether or not they violate the rule.	sealed	unsealed	50 lire	40 lire

Core concepts

People make some conditional inferences very readily, and they make mistakes with others. Their mistakes provide clues about the sorts of mental representations and cognitive processes that they rely on to make inferences.

Conditional inferences

Suppose you are told that if there was a lightning storm last night, there was a fire in the old forest. Suppose you then discover that there was indeed a lightning storm last night. You will find it easy to make the inference that there was a fire in the old forest. Usually close to 100% of participants make this simple *modus ponens* inference in the many experiments that have examined it (for a review, see [Nickerson, 2015](#)). It has the form, “If A, then B. A, therefore B” (see [Table 2](#)). It is considered valid within propositional logic because the conclusion must be true when the premises are true, that is, when the information on which the conclusion is based is true. Now suppose instead that what you discover is that there was no fire in the forest. Should you make the inference that there was no lightning storm last night? The inference seems more difficult than the previous one. About 60% to 70% of participants make this *modus tollens* inference, but about 30% to 40% say that nothing follows. It has the form, “If A, then B. Not-B, therefore not-A.” The inference is also valid according to propositional logic, but it seems less obvious than the previous one, and there seems to be less agreement among people about whether it is a good inference to make.

Consider another inference, this time one that is often considered invalid according to propositional logic. Suppose you discover that there was a fire in the old forest. Was there a lightning storm last night? You may be tempted to make the *affirmation of the consequent* inference, that there was a lightning storm, and about 70% of participants do so (e.g., [Nickerson, 2015](#)). It has the form, “If A, then B. B, therefore A.” However, there are other causes of forest fires, such as campfires that get out of hand, and so the conclusion that there was a lightning storm is not necessarily true. The inference is considered a fallacy by the normative standard of propositional logic. For the same reason, it is also considered a fallacy to make the *denial of the antecedent* inference from the information that there was no lightning storm to the conclusion that there was no fire. About 60% of participants in experiments make this inference, which is of the form, “If A, then B. Not-A, therefore not-B.”

The suppression of inferences

An important clue about the cognitive processes that people rely on to make conditional inferences comes from the observation that they sometimes refrain from making any of them. For example, when participants are told that if there was a lightning storm last night, there was a fire in the old forest, and they are also told about other alternative causes, such as if campers lit a campfire, there was a fire in the old forest, they do not make the *affirmation of the consequent* inference or the *denial of the antecedent* one. They know there may have been a campfire instead. The alternative cause suppresses the fallacies.

Strikingly, when participants are told that if there was a lightning storm last night, there was a fire in the old forest, and they are also told about other enabling causes, such as if there were dry leaves on the forest floor, there was a fire in the old forest, they do not make the *modus ponens* inference or the *modus tollens* one. They know there must have been dry leaves too. The enabling cause suppresses the valid inferences. Even the simplest inference is suppressed when background knowledge conflicts with it ([Byrne, 1989](#); see [Espino & Byrne, 2020](#) for a review).

Table 2

The Form of Four Conditional Inferences

Table 2		
Inference	Premises	Conclusion
<i>Modus ponens</i>	If A, then B. A.	Therefore B.
<i>Modus tollens</i>	If A, then B. Not-B.	Therefore not-A.
<i>Affirmation of the consequent</i>	If A, then B. B.	Therefore A.
<i>Denial of the antecedent</i>	If A, then B. Not-A.	Therefore not-B.

Questions, controversies, and new developments

There is currently no consensus in cognitive science about the mental representations and cognitive processes that explain deductive reasoning. Nor is there agreement about whether the accuracy of human inferences can be best judged by comparison to formal logics or to other normative accounts. Instead, there are several prominent theories of how people reason.

One of the earliest psychological theories of human deductive reasoning proposed that it depends on a mental logic consisting of abstract inference rules (e.g., [Braine & O'Brien, 1998](#); [Rips, 1994](#)). On this account, when people make an inference, they first strip away the content of the premises to uncover their logical skeleton. For example, the underlying form of the conditional premise, if there was a lightning storm last night, there was a fire in the old forest, and the additional premise, there was a lightning storm last night, is, “If A, then B. A.” People match this form to one of the elementary inference rules in their mental logic, such as one corresponding to *modus ponens*, “If A, then B. A, therefore B,” and so they can conclude B. Their mental logic does not contain a rule corresponding to the *modus tollens* inference, and so instead they have to construct a mental derivation of the conclusion using other elementary rules, akin to constructing a proof in logic. The theory assumes the accuracy of human inferences can be assessed by comparison to formal logics.

Inference rule theories of this sort were criticized because they could not readily explain why inferences are affected by content, such as in the selection task or the suppression task (see [Nickerson, 2015](#)). Accordingly, another early theory of reasoning proposed that the mind contains content-sensitive inference rules such as, “If the action is to be taken, the precondition must be met,” or, “If the benefit is taken, the cost must be paid” (e.g., [Cheng & Holyoak, 1985](#); [Cosmides, 1989](#)). Domain-specific inference rule theories have been criticized for being unable to explain how people can make inferences in situations that do not concern preconditions and actions or costs and benefits (e.g., [Sperber et al., 1995](#)).

A very different theory of how people make deductive inferences proposes that people construct mental models ([Johnson-Laird, 1983](#)). Unlike inference rules, the mental model theory proposes that people reason by constructing small-scale iconic models of information about real or imagined situations ([Johnson-Laird & Byrne, 1991](#)). When people understand that if there was a lightning storm last night, there was a fire in the old forest, they construct an initial simulation of a single possibility corresponding to the way the world would be if the assertion were true: there was a lightning storm last night, and there was a fire in the old forest. On reflection, they can also think about other possibilities consistent with the assertion, such as that there was no lightning storm last night, and there was no fire in the old forest, and even, there was no lightning storm last night, and there was a fire in the old forest anyway. They can come to many different interpretations of conditionals ([Johnson-Laird & Byrne, 2002](#)). They can make the *modus ponens* inference readily because the information that there was a lightning storm is consistent with the initial possibility they thought about, and they can conclude there was a fire. The information that there was no fire is not consistent with the initial possibility, and so some participants may consider nothing

follows; the *modus tollens* inference requires them to consider the other possibilities consistent with the assertion. The theory proposes that the accuracy of human inferences can be assessed by whether a conclusion is a necessary one for which there are no counterexamples ([Khemlani, Byrne, & Johnson-Laird, 2025](#)).

The model theory is a prominent contemporary theory of human reasoning (see <https://www.modeltheory.org/>). It has been criticized for proposing that possibilities, not probabilities, are the basis for human reasoning (e.g., [Evans & Over, 2004](#); [Oaksford & Chater, 2007](#)). An alternative probabilist theory proposes that people understand an assertion—such as if there was a lightning storm last night, there was a fire in the old forest—by calculating their degrees of belief in the probability that there was a fire in the old forest, given that there was a lightning storm last night [see [Foundations of Rationality](#); [Human Rationality in Research](#)]. When they discover there was a lightning storm last night, they make the inference that there was a fire in the old forest only if their degrees of belief in the conditional probability are high. It proposes that the accuracy of human inferential conclusions should be judged by comparison to the normative standards of Bayesian probability [see [Bayesian Models of Cognition](#)]. It has been criticized for failing to specify the nature of the mental representations and cognitive processes that people rely on to make inferences (e.g., [Knauff & Gazzo Castaneda, 2023](#)).

Current research tests the predictions of these alternative theories, and many contemporary experiments pit the predictions of the model theory against probabilist theories. Current experiments often examine more complex conditional inferences as well as the relation between conditional inferences and inferences from other propositional connectives such as disjunctions. They test the mixture of fast intuitive processes and slower reflective processes that deductive inferences depend on, similar to other sorts of thinking (e.g., [De Neys, 2023](#)).

An important development in contemporary studies of hypothetical thought is the study of people's ability to make hypothetical inferences about things that did not happen. Many experiments examine inferences from counterfactual conditionals, such as if there hadn't been a lightning storm, there wouldn't have been a fire in the old forest. People make inferences more readily from counterfactuals than other conditionals. For example, when participants discover that in fact there was no fire in the old forest, they can make the *modus tollens* inference that there was no lightning storm far more readily than from a factual conditional (see [Byrne, 2016](#) for a review). Counterfactual reasoning is central not only to causal understanding but also to moral cognition (e.g., [Byrne, 2005](#); [Gerstenberg et al., 2017](#)).

Broader connections

Hypothetical deductive reasoning is closely linked to other sorts of reasoning including abductive reasoning, analogy, and hypothesis testing [see [Analogy](#)]. It is highly relevant to moral cognition and to modal reasoning [see [Morality](#)]. It is a key aspect of human thinking, and its study addresses similar issues that arise in the study of decision-making and problem-solving, including the role of heuristics. The experimental study of deduction in psychology is closely linked to developments in the study of

logics in philosophy, inference defeasibility and presuppositions in linguistics, and non-monotonicity and belief revision in artificial intelligence.

Acknowledgments

The author would like to thank Mark Keane and Phil Johnson-Laird for helpful discussion.

Further reading

- Ball, L. J., & Thompson, V. A. (Eds.). (2017). *International handbook of thinking and reasoning*. Routledge.
- Byrne, R.M.J (2005). *The rational imagination: How people create alternatives to reality*. MIT press.
- Knauff, M., & Spohn, W. (Eds.). (2021). *The handbook of rationality*. MIT Press.
- Nickerson, R. (2015). *Conditional reasoning: The unruly syntactics, semantics, thematics, and pragmatics of “if.”* Oxford University Press.

References

- Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Lawrence Erlbaum.

↩

- Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31(1), 61-83.
[https://doi.org/10.1016/0010-0277\(89\)90018-8](https://doi.org/10.1016/0010-0277(89)90018-8)

↩

- Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. MIT Press.

↩

- Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, 61, 135-157.
<https://doi.org/10.1146/annurev-psych-122414-033249>

↩

- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17(4), 391-416.
[https://doi.org/10.1016/0010-0285\(85\)90014-3](https://doi.org/10.1016/0010-0285(85)90014-3)

↩

- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31(3), 187-276. [https://doi.org/10.1016/0010-0277\(89\)90023-1](https://doi.org/10.1016/0010-0277(89)90023-1)

↩

- De Neys, W. (2023). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences*, 46, e111. <https://doi.org/10.1017/S0140525X2200142X>

↩

- Espino, O., & Byrne, R. M. (2020). The suppression of inferences from counterfactual conditionals. *Cognitive Science*, 44(4), e12827. <https://doi.org/10.1111/cogs.12827>

↵

- Evans, J. S. B. T., & Over, D. (2004). *If*. Oxford University Press.

↵

- Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, 28(12), 1731-1744. <https://doi.org/10.1177/0956797617713053>

↵

- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge University Press.

↵

- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Erlbaum.

↵

- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological review*, 109(4), 646-678. <https://doi.org/10.1037/0033-295X.109.4.646>

↵

- Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive science*, 42(6), 1887-1924. <https://doi.org/10.1111/cogs.12634>

↵

- Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427-457. <https://doi.org/10.1037/a0026841>

↵

- Knauff, M., & Gazzo Castaneda, L. E. (2023). When nomenclature matters: Is the “new paradigm” really a new paradigm for the psychology of reasoning? *Thinking & Reasoning*, 29(3), 341-370. <https://doi.org/10.1080/13546783.2021.1990126>

↵

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

↵

- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

↵

- Ragni, M., Kola, I., & Johnson-Laird, P. N. (2018). On selecting evidence to test hypotheses: A theory of selection tasks. *Psychological Bulletin*, 144(8), 779-796. <https://doi.org/10.1037/bul0000146>

[↩](#)

- Rips, L. (1994). *The psychology of proof*. MIT Press.

[↩](#)

- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57(1), 31-95. [https://doi.org/10.1016/0010-0277\(95\)00666-M](https://doi.org/10.1016/0010-0277(95)00666-M)

[↩](#)