Forthcoming in *The Encyclopedia of the Mind*, edited by Hal Pashler, SAGE Publishing.
Editorial Board: Tim Crane, Fernanda Ferreira, Marcel Kinsbourne, and Rich Zemel.

Deductive Reasoning

Joshua Schechter
Brown University

Deductive reasoning is the kind of reasoning in which, roughly, the truth of the input propositions (the premises) logically guarantees the truth of the output proposition (the conclusion), provided that no mistake has been made in the reasoning. The premises may be propositions that the reasoner believes or assumptions that the reasoner is exploring. Deductive reasoning contrasts with inductive reasoning, the kind of reasoning in which the truth of the premises need not guarantee the truth of the conclusion.

For example, a reasoner who infers from the beliefs

(i) If the room is dark then either the light switch is turned off or the bulb has burned out;

(ii) The room is dark;

(iii) The light switch is not turned off;

to the conclusion

(iv) The bulb has burned out;

is reasoning deductively. If the three premises are true, the conclusion is guaranteed to be true. By contrast, a reasoner who infers from the belief

(i) All swans that have been observed are white;

to the conclusion

(ii) All swans are white;

is reasoning inductively. The premise provides evidential support for the conclusion, but does not guarantee its truth. It is compatible with the premise that there is an unobserved black swan.

Deductive reasoning has been intensively studied in cognitive science, psychology, and philosophy. There are many important debates concerning the nature of deductive reasoning. This entry surveys three topics – the relationship between deductive reasoning and logic, the main psychological models of deductive reasoning, and the epistemology of deductive reasoning.

Deductive Reasoning and Logic

Deductive reasoning should be distinguished from logic. Deductive reasoning is a psychological process. In contrast, logic does not describe a psychological process. Logic is the abstract theory of the logical consequence relation, the relation that specifies what follows from what.

While logic does not provide a descriptive theory of reasoning, it is widely accepted that there is a normative connection between logic and reasoning. Yet, it has proved difficult to precisely articulate this connection. On a simple proposal, reasoners ought to infer the logical consequences of their beliefs and ought not to hold logically inconsistent beliefs. However, Gilbert Harman and others have provided several arguments against this proposal. For example, in some cases, when reasoners recognize that an implausible claim follows from their beliefs, they should not believe the claim but instead give up one of their antecedent beliefs. Harman further claims that there is no distinctive tie between logic and reasoning. This claim, however, is not widely endorsed.

Despite the difficulties of precisely articulating the connection between logic and reasoning, the idea that logic provides a normative standard for reasoning is widely accepted.

The Cognitive Psychology of Deductive Reasoning

Sources of Evidence

The psychological study of deductive reasoning has largely focused on investigating its algorithmic underpinnings. Experimental evidence has come from the performance of subjects, typically undergraduate students, on specific reasoning tasks. In one experimental paradigm, subjects are presented with premises and asked to indicate whether a proposed conclusion follows. In a related paradigm, subjects are presented with premises and asked to generate a conclusion that follows, if one exists.

A notable finding of this research is that subjects are highly prone to errors in evaluating the validity of arguments. For example, a meta-analysis carried out by Walter Schroyens, Walter Schacken, and Géry D'Ydewalle of 65 studies on conditional reasoning found that while 97% of subjects correctly evaluated Modus Ponens inferences (if p then q, p, therefore q) as valid, only 72% correctly evaluated Modus Tollens inferences (if p then q, not-q, therefore not-p) as valid. Worse still, 63% incorrectly identified instances of the fallacy of Affirming the Consequent (if p then q, q, therefore p) as valid and 55% incorrectly identified instances of the fallacy of Denying the Antecedent (if p then q, not-p, therefore not-q) as valid. This research has also uncovered several content effects that influence subjects' responses. For example, subjects are more likely to mistakenly identify an invalid argument as valid if its conclusion is believable.

Another important source of evidence comes from the Wason selection task, a paradigm developed by Peter Wason. In a typical experiment, subjects are shown four cards and told that each card has a letter on one side and a numeral on the other. The visible sides of the cards show D, K, 3, and 7, respectively. Subjects are presented with the conditional claim: Every card which has a D on one side has a 3 on the other side. Subjects are then asked which cards they need to turn over to determine whether the conditional is true. The correct answer is the D and 7 cards, since the only way to falsify the conditional is for a card to have a D on one side without a 3 on the other. Very few subjects (typically around 10%) select these two cards.

Interestingly, subjects perform much better on variants of this task. In an experiment carried out by Richard Griggs and James Cox, subjects are told to imagine that they are police officers observing people in a bar. They are shown four cards each representing a person at a table, with the person's age on one side and the person's beverage on the other. The visible sides show "drinking a beer," "drinking a coke," "16 years of age," and "22 years of age." Subjects are presented with the rule: If a person is drinking beer, then the person must be over 19 years of age. Subjects are then asked which cards they need to turn over to determine whether the rule is being violated. Most subjects (74%) correctly select the "drinking a beer" and "16 years of age" cards.

Many variants of the Wason selection task have been investigated. Subjects do better in certain variants than others. It is controversial how these cases should be distinguished. Apparently minor changes in wording can have a significant impact. To a first approximation, however, subjects perform better on variants involving "realistic"

rather than abstract contents and on variants that involve permissions or obligations rather than descriptive conditionals.

The data concerning errors made in deductive reasoning provide the principal way of testing psychological theories of deductive reasoning.

Mental Logic

There are two main rival psychological theories of deductive reasoning – the mental logic theory and the mental models theory. Versions of the mental logic theory have been championed by Martin Braine and David O'Brien and by Lance Rips, among others. The central claims of this theory are as follows: Human reasoning makes use of mental representations that resemble the sentences of natural language. In deductive reasoning, reasoners manipulate these representations by applying syntactic rules of inference that resemble the rules of logic.

Versions of the mental logic theory differ over exactly which rules are employed in deductive reasoning. Typically, however, they claim that the rules resemble the rules that appear in natural deduction formulations of formal logic. Some of these rules involve the use of suppositions. For example, the rule of Reductio ad Absurdum states that if supposing that p leads to an absurdity, then one can conclude that not-p.

Mental logic theories explain errors in deductive reasoning tasks by appealing both to the difficulty of applying particular rules and to the need to apply multiple rules in particular tasks. For instance, such theories typically claim that it is more difficult to identify Modus Tollens as valid (compared to Modus Ponens) because there is no mental rule corresponding to Modus Tollens. Reasoners must rely on multiple rules (Modus

Ponens and Reductio) to derive its validity. Mental logic theories do not directly explain the effects of content on reasoning. Such effects are typically explained by appeal to additional cognitive processes.

Mental Models

The mental models theory has been championed by Philip Johnson-Laird and his collaborators. On this view, deductive reasoning involves diagrammatic rather than language-like representations. In an inference, reasoners construct models that represent the possible states of the world compatible with the premises. They then formulate a putative conclusion that is relevant and informative. Finally, they test this conclusion to make sure that the models do not provide a counterexample to it.

For example, the conditional premise, if p then q, can be exhaustively represented with the following three models, each specifying a possible state of the world consistent with the premise:

p q

not-p q

not-p not-q

The mental models theory claims that, to save working memory, reasoners do not typically exhaustively represent the information provided by the premises. An initial mental representation of the conditional might instead be:

[p] q

…

The square brackets indicate that the first model is the only model in which p is true. The ellipsis indicates that there may be additional models beyond those explicitly represented.

In a Modus Ponens inference, the initial mental representation of the premise that if p then q is combined with the representation of the premise that p. This eliminates the implicit non-p models represented by the ellipsis. The result is the single model:

p q

Since this is a model in which q is true, there are no models that provide counterexamples to q. This yields the information that q follows from the premises.

The mental models theory explains errors in deductive reasoning tasks by appealing to the difficulty of reasoning with multiple models. For example, it is more difficult to identify Modus Tollens as valid (compared to Modus Ponens) because Modus Tollens cannot be shown to be valid using the initial mental representation of the conditional premise. Rather, it requires "fleshing out" the initial mental representation of the conditional premise into the exhaustive representation, and combining this with the information that not-q. The combination eliminates the first two models, leaving the single model:

not-p not-q

This is a model of not-p, and so yields the information that not-p follows from the premises. This process is more complicated than the process for Modus Ponens, thus explaining the greater difficulty of identifying Modus Tollens as valid.

The mental model theory can also explain certain content effects. For example, the bias in favor of believable conclusions is explained by the lack of motivation subjects have to search for counterexamples to believable claims.

Alternative Views

While the mental rules and mental models theories have many differences, they both claim that the cognitive mechanism for deductive reasoning is a general-purpose reasoning mechanism central to reasoning and problem solving. There are several alternative views that deny this claim. One view is that humans do not possess a general-purpose mechanism for deductive reasoning, but rather a different kind of general-purpose reasoning mechanism, for example one devoted to probabilistic or explanatory reasoning.

A different view is that humans lack general-purpose reasoning mechanisms (or that such mechanisms are not central to reasoning) but instead employ many special-purpose reasoning mechanisms. This view is often motivated by the performance of subjects in the Wason selection task. The idea is that the greater success of subjects in certain variants of the task is explained by the presence of a special-purpose reasoning mechanism. Patricia Cheng and Keith Holyoak have argued that humans make use of pragmatic reasoning schemas, sets of learned context-sensitive rules tied to particular goals. For example, they argue that we possess rules for reasoning about permissions and obligations. Leda Cosmides has argued that humans possess an innate special-purpose cognitive module for detecting cheating in social exchanges. On both of these views, general-purpose reasoning mechanisms do not play an important role in cognition.

Dual Process Theories

A common view among psychologists is that humans have two reasoning systems. System 1 is a relatively fast, cognitively undemanding, automatic reasoning system.

System 2 is a relatively slow, cognitively demanding, deliberate reasoning system. Versions of this theory differ in how the two systems should be distinguished and how they interact with one another. The cognitive mechanism for deductive reasoning is presumably a part of System 2.

The Epistemology of Deductive Reasoning

A central topic in epistemology concerns the nature of justified belief (roughly, reasonable or rational belief). Competently performed deductive inferences typically preserve justification. For example, if a thinker forms a belief via a Modus Ponens inference applied to justified beliefs, the resulting belief will typically also be justified. This raises the question of why competent deductive reasoning preserves justification. This is not a skeptical worry about whether justification is preserved but rather an explanatory question about why deductive inference has this positive normative status.

There are several different theories of why deductive reasoning preserves justification. On one view, the explanation is that competent deduction is a reliable cognitive process. If the premises are true, the conclusion is guaranteed also to be true. This is a version of reliabilism, a view developed by Alvin Goldman. Reliabilism states that a thinker is justified in holding a belief if it is the output of a cognitive process that either (i) does not depend on any input beliefs and tends to yield truths or (ii) depends on input beliefs and tends to yield truths when the input beliefs are true.

Reliabilism faces several problems. For instance, it has difficulty explaining why thinkers who are unreliable through no fault of their own – perhaps because they are being fed misleading experiences by René Descartes' imagined evil demon – can

nevertheless be justified in many of their beliefs. A second problem is that there are reliable deductive inferences that do not preserve justification. For example, inferring the conclusion of a complicated proof directly from its premises would not yield a justified belief. To have a justified belief, one must be aware of the proof. These examples suggest that reliabilism has difficulty capturing the intuitive connection between having a justified belief and being a responsible thinker.

A second view is that a pattern of reasoning only preserves justification if the relevant thinker has an appreciation of the fact that the pattern is truth-conducive. Versions of this view differ on the nature of this appreciation. On one version, thinkers have an empirically justified belief that the pattern preserves truth. On another version, thinkers instead have an a priori justified belief. (This belief may be generated by an a priori faculty of intuition or rational insight.) One difficulty facing this view is that young children can have justified beliefs arrived at by deduction even though they are not sophisticated enough to appreciate that their pattern of reasoning is truth-conducive. A second difficulty is that this view is prone to regress worries. For example, if the belief that deductive reasoning is truth-conducive plays an inferential role, presumably the thinker also must appreciate the fact that this more complex inference is truth-conducive, thus launching a regress.

A third view connects justification with the nature of concepts. Versions of this view have been endorsed by Paul Boghossian and Christopher Peacocke, among others. This view involves two main ideas. The first idea is that to possess certain logical concepts, thinkers must employ particular deductive rules. For example, to possess the concept of the conditional, thinkers must employ the rule Modus Ponens. The second

idea is that any rule that is constitutive of a genuine logical concept preserves justification. This view also faces several difficulties. The claim that possessing logical concepts requires employing particular deductive rules is contentious. Putting this worry aside, it is plausible that not every concept has justification-preserving constitutive rules. It is difficult to provide a principled distinction between concepts that have justification-preserving constitutive rules and those that do not. Finally, this view does not easily generalize to explain why inductive reasoning can preserve justification. Since there is presumably a unified account of inferential justification, this poses a problem for the view.

There are many other views of epistemic justification. The correct account of the epistemology of deductive reasoning is a matter of ongoing research.

See also Inductive Reasoning; Knowledge, Philosophical Perspectives; Logic and Its Relation to Language; Thinking

Further Readings

BonJour, L. (1998). *In defense of pure reason*. Cambridge: Cambridge University Press.

Boghossian, P. (2003). Blind reasoning. *Proceedings of the Aristotelian Society, Supplementary Volume*, 77, 225–248.

Evans, J., Newstead, S. & Byrne, R. (1993). *Human reasoning.* Hillsdale: Lawrence Earlbaum.

Harman, G. (1986). *Change in view: Principles of reasoning*. Cambridge, MA: The MIT Press.

Johnson-Laird, P. & Byrne, R. (1991). *Deduction*. Hillsdale: Lawrence Earlbaum.

Rips, L. (1994). *The psychology of proof*. Cambridge, MA: The MIT Press.