Chapter 7

Deductive Reasoning

JONATHAN ST. B. T. EVANS

University of Plymouth

A deduction is a conclusion that follows from things we believe or assume. Frequently, we combine some fact or observation with a rule or rules that we already believe. For example, you meet Sue, who tells you that Mary is her mother. You immediately infer that Sue is Mary's daughter, although that is not the fact that was presented to you. This means that you must carry around rules such as

> If x is the mother of y and y is female, then y is the daughter of x

Of course, I am not saying that you are conscious of this rule or of applying it to make this deduction but it must be there in some form in your brain, together with some mechanism for applying it to facts and observations. In this case, the inference occurs rapidly and effortlessly but this is not always the case with deductive reasoning. Take the case of claiming allowances when completing an income tax return. In this case there may be many rules and their wording may be complex and opaque to those who are not expert in tax law. If your financial affairs are complex, even establishing the relevant facts may be headache. This is why people often pay expert tax advisers to do the reasoning for them.

Deduction has a clear and obvious benefit for human beings. Our memories are limited and our brains can only store so many beliefs about the world. However, if we also hold a number of general rules, then these can be applied to draw out implications as and when they are required. Deduction is also involved in hypothetical thinking, when we ask 'What if?' questions. An example is science in which theories must be tested against empirical evidence. Scientific theories take the form of rules, often formalised with mathematics. When experimental studies are run, we set up some conditions and then predict the outcome. The prediction is a deduction, which is used to test the theory. For example, climate change scientists have been predicting for the past twenty years or more that warming temperatures would disrupt the jet stream and lead to more extreme weather events. These predictions were calculated from their mathematical models, which is a form of deductive reasoning. Both abnormal jet stream flows and extreme weather events have been observed in recent years with increasing frequency, lending credibility to these models.

For deduction to be useful, it needs to be accurate. This has been recognised in the discipline of philosophy for centuries. Philosophers devised systems of *logic*, whose purpose is to ensure accurate deduction. A logically valid argument is one whose conclusion necessarily follows from its premises. Put simply, this means that in a logical argument the conclusion must be true if the premises are true. If the mother-daughter rule given earlier is true (which it is by convention) and your observation that Sue is female is also correct, then she *must* be Mary's daughter.

Logic provides rules for reasoning. Here are a couple of examples

https://doi.org/10.17885/heiup.470.c6670

Modus Ponens Given if x then y, and the assumption of x, y is a valid conclusion

Disjunction elimination Given x or y and not-x, y is a valid conclusion

Modus Ponens is very useful, because it means we can state hypothetical beliefs, which only apply when some condition is met. Some of the conditional sentences we use in everyday life are necessarily true, for example, 'if a number is even and greater than two, it cannot be prime', but most are not. For example, we may advise someone 'if you catch the 8.00 am train then you will get to work on time'. If this is generally true, then it is good advice, but of course the train might break down. The real world rarely allows inferences to be certain, but we nevertheless use conditional statements a great deal because of the natural power of Modus Ponens. A disjunctive statement is an either-or. For example, someone might say 'I will either catch the 8.00 am train or take the bus at 8.10'. If you later learn that they did not catch the train you can deduce that they took the bus instead. Once again, in the real world, the inference will not be certain. The individual may have called in sick and not gone to work at all. But our deduction is valid, given the assumptions on which is based.

There is a tradition in philosophy that logic is the basis for rational thought (Henle, 1962; Wason & Johnson-Laird, 1972). This view held a powerful influence on psychology during the second half of the twentieth century and was responsible for a major method of studying human reasoning, which I will call the **deduction paradigm** (Evans, 2002). Huge numbers of experiments were run with this paradigm and I will try to summarise their main findings in this chapter. While many important things were learnt about the nature of human thinking and reasoning, a lot of psychologists eventually lost faith in the importance of logic for rational thinking. In recent years, this has a led many to revise their methods and adopt what is called the **new paradigm** psychology of reasoning. I will explain the new paradigm and some of the findings it has led to at the end of this chapter. For now, I will focus on the deduction paradigm and the theories and findings that are associated with it.

The deduction paradigm tests whether people *untrained in logic* can make valid inferences. The idea behind this is that if logic is the basis for rationality in everyday thinking then everyone should comply with it, not just those who have taken logic classes. So the first condition in this method is to exclude participants with formal training. The next is to present them with some premises, or assumptions, from which a logical deduction can be made. Often a conclusion is also given and people are asked whether it follows or not. Two other instructions are usually given: (1) assume that the premises given are true and (2) only make or endorse a conclusion which necessarily follows from them. Given these instructions, only the form of the argument should matter, not the content. For example, people should always agree that Modus Ponens and disjunction elimination are valid arguments, no matter what we substitute for x and y in the rules given above. By this means, the paradigm assesses whether or not people are logical in their deductive reasoning.

7.1 The Deduction Paradigm: The Main Methods and Findings

A small number of experiments on deductive reasoning were published early in the twentieth century (Wilkins, 1928; Woodworth & Sells, 1935), which immediately demonstrated what was to come from the intensive study that occurred from the 1960s onwards. That is to say, people were observed to make frequent logical errors, to show systematic biases, and to be influenced by their beliefs about the content of the premises and conclusions. All of these findings have been replicated many times since using three major methods, or sub-paradigms. These are syllogistic reasoning, conditional inference, and the Wason selection task. In this section I will discuss each in turn, explaining the methods and typical findings.

7.1.1 Syllogistic Reasoning

A syllogism involves two premises and three terms, which I will call A, B and C. This is the most ancient system of logic, devised by Aristotle. You may have come across the famous syllogism 'All men are mortal, Socrates is a man, therefore Socrates is mortal.' Classical syllogisms have statements in four *moods*, shown in Table 7.1 (a). These statements can be used for either the first premise, the second premise, or the conclusion in any combination. Let us consider them in turn.

Figure 7.1 shows diagrammatically several different models for the relation between two categories, A and B. When we examine the different statements in Table 7.1 (a) we see that most of them are ambiguous and can be represented by at least two different models. For example, All A are B would be true for a model of identity – All men have Y chromosomes – or where B includes A – All boys are male. No A are B is unambiguous; it can only refer to a model of exclusion. Some A are B is highly ambiguous – it is true in all models except exclusion. Finally, Some A are not B is true for exclusion but also for a model in which A includes B – Some males are not boys. This gives us a clue to the complexity of syllogistic reasoning, as we have to take account of all possible ways that the categories could be related. Moreover, when we combine two premises, we have to consider the ways in which all three categories A, B, and C could be related. For the argument to be valid, its conclusion has to be true in all models of this three-way relationship that the premises allow.

A fallacy is an argument whose conclusion need not be true, given the premises. A basic finding with syllogistic reasoning is that participants endorse many fallacies. This has been reported by many authors and confirmed in the one study (to my knowledge) that presented every possible combination of premises and conclusions for evaluation (Evans, Handley, Harper, & Johnson-Laird, 1999). But these mistakes are not random – there are systematic biases in syllogistic reasoning. Consider the following syllogism:

All A are B

All B are C

Therefore, All C are A

This is a fallacy: the conclusion does not necessarily follow. Yet in the study of Evans et al. (1999), 77% of participants (university students) said that the conclusion necessarily followed from

Table 7.1: The structure of classical syllogisms.

(a) Mood of premises

A	All A are B
E	No A are B
I	Some A are B
O	Some A are not B

(b) Figure of syllogism

1	2	3	4
A – B	A – B	B – A	B – A
B – C	C – B	B – C	C – B
——	——	——	——
A – C	A – C	A – C	A – C

Note: The letters A, E, I and O are classically used as abbreviations for the four moods of the premises.

the premises. This is really odd when you consider the set relationships involved (see Figure 7.2). Surely, the most likely situation that describes the premises is the model to the left, showing that A is a subset of B and B a subset of C. For example

All Alsatians are dogs; all dogs are mammals

But in that the case, the conclusions endorsed would be 'All mammals are Alsatians,' which is obviously false. The conclusion All C are A would only be true in the second model, where A, B, and C are all identical, a most unusual state of affairs. This is a finding which you only get with abstract materials, where letters are used to represent categories. But why does it occur? It is consistent with a very old claim called the *atmosphere effect*: participants are inclined to accept conclusions whose mood matches that of the premises. In the same study, only 47% of participants said the following syllogism was valid

All A are B

All B are C

Therefore, Some C are A

This is stranger still because Some C are A has to be true whenever All C are A. Not only that, but the conclusion is actually *valid* in this case. You can verify that by examining the models of the premises shown in Figure 7.2. Of course, the mood of the conclusion does not match that of the premises here, so it does conform with the atmosphere effect. In fact, atmosphere is consistent with many but not all responses observed in syllogistic reasoning tasks (Evans, Newstead, & Byrne, 1993). Another known biasing factor is the figure of the syllogism, which is the order in which terms are arranged (Table 7.1b). This also affects people's perception of validity.

There are a number of high-profile theories of syllogistic reasoning based on different principles and giving broadly accurate explanations of the data (for reviews, see Evans et al., 1993; Manktelow, 1999). When realistic content is introduced, however, other factors come into play, especially **belief bias**. Consider the following syllogism:

No addictive things are inexpensive

Some cigarettes are inexpensive

Therefore, some addictive things are not cigarettes 71%

A major study which established the influence of beliefs was that of Evans, Barston, and Pollard (1983). Over three experiments, they found that

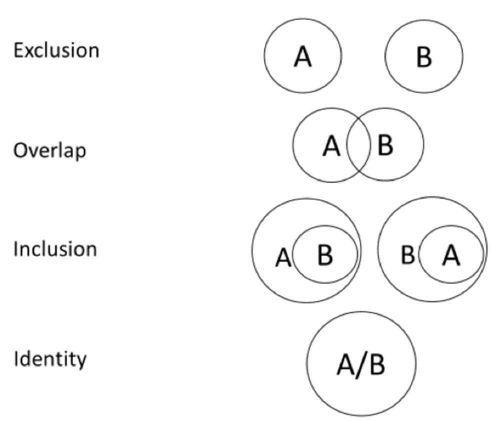

Figure 7.1: Models of relations between two categories, A and B.

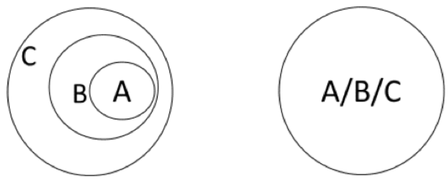

Figure 7.2: Models of relations between three categories, A and B and C consistent with the premises All A are B, All B are C.

71% of participants agreed that this syllogism was valid, that is, that the conclusion must be true if the premises are true. Now consider this syllogism presented in the same study:

No millionaires are hard workers

Some rich people are hard workers

Therefore, some millionaires are not rich people 10%

In this case, only 10% thought the syllogism was valid. But if you look carefully you can see that both syllogisms have the same logical form. The syllogism is actually *invalid*. The key difference between these two is that the first realistic version has a conclusion which is believable (we know that there are other addictive drugs) but the second has a conclusion which is unbelievable – millionaires are rich by definition. When valid arguments were used, people also more often thought they were valid if they believed the conclusion (89%) than if they did not (56%), so there is a belief bias for valid arguments as well, although not as a strong. These findings have been replicated many times since (for a review see Klauer, Musch, & Naumer, 2000).

7.1.2 Conditional Inference

Conditional statements, also known just as conditionals, have the form, if p then q. We use many such statements in real life for all kinds of purposes. Here are some examples:

Causal: If you heat water sufficiently, it will boil.

Prediction: If you vote Republican, you will get your taxes cut.

Tip: If you study hard, you will pass the examination.

Warning: If you miss the 8.00 am train, you will be late for work.

Promise: If you wash my car, I will give you ten dollars.

Threat: If you stay out late again, you will be grounded.

Counterfactual: If you had putted well, you would have won the match.

Philosophers have long studied conditional statements, considering them of particular importance for human reasoning, writing many books on the subject (an excellent review of philosophical work

Table 7.2: The four main conditional inferences.

Label	Rule	Example
Modus Ponens (MP)	If p then q; p: therefore q.	If the letter is B then the number is 3; the letter is B: therefore the number is 3.
Denial of the Antecedent (DA)	If p then q; not-p; therefore not-q.	If the letter is G then the number is 7; the letter is not G: therefore, the number is not 7.
Affirmation of the Consequent (AC)	If p then q; q: therefore p.	If the letter is T then the number is 5; the number is 5: therefore the letter is T.
Modus Tollens (MT)	If p then q; not-q; therefore not-p.	If the letter is M then the number is 1; the number is not 1: therefore, the letter is not M.

is given by Edgington, 1995). A great deal of work in the psychology of reasoning has also focussed on conditional statements (for a recent review see Nickerson, 2015). I wrote an entire book on 'if' myself, collaborating with a philosopher to cover the perspectives of both traditions—philosophy and psychology (Evans & Over, 2004). The reason they are so important is that they are central to a unique human facility which I call *hypothetical thinking* (Evans, 2007). That is the ability to imagine how things might be in the future or how they might have been different in the past.

Standard logic provides an account of how we should reason with conditionals. Some of this standard account is disputed by both philosophers and psychologists, but all are agreed about the four inferences shown in Table 7.2. We have already encountered Modus Ponens (MP) as an example of a valid deductive argument. Imagine a situation where cards have a letter written on one side and a number on the other. Then we can express a conditional hypothesis such as

If the letter is B, then the number is 3

For conditional inference, we need to assume that the conditional is true. This is the *major premise* of the deductive argument. The *minor premise* is an assertion that either the first or second part of the conditional is true or false, leading to the four arguments illustrated in Table 7.2. For example, if we suppose that the first part is true – the letter is B – then it follows by MP (Modus Ponens) that the number is 3. The other valid argument that can be made is Modus Tollens (MT). Suppose that the number on the card is not a 3. Then it follows logically that the letter is not a B. Why? Because the conditional is true, so if there had been a B on the card, then there would also have to have been a 3. This argument is equally valid, if less immediately obvious. What all psychological studies of such abstract inferences show is that the MP inference is made nearly 100% of the time, while the MT inference is only endorsed about 60% of the time in the same experiments (see Figure 7.3).

The other two arguments shown in Table 7.2 are fallacies. That is, the conclusions given do not *necessarily* follow. If we assume that the number is 3, it does not necessarily follow that the letter is a B (Affirmation of the Consequent, AC), because the conditional does not say that only B cards can have 3s. Similarly, if we know that the letter is not B, we cannot say that the number is not a 3 (Denial of the Antecedent, DA). And yet we see that university students endorse both of these fallacies about 40% of the time (see Figure 7.2). A likely reason for this is that people are making inferences that are *pragmatically* rather than logically implied. For example, the Denial of the Antecedent fallacy

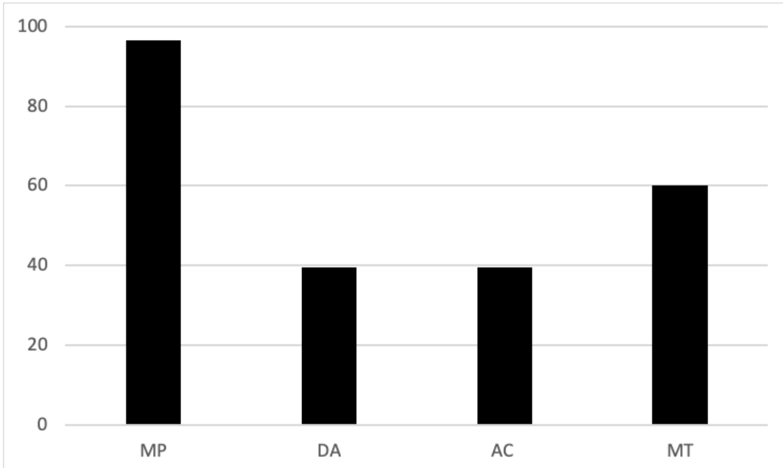

Figure 7.3: Percentage frequencies of endorsement of the four conditional inferences with abstract materials (weighted average of 11 studies reported by Evans et al. 1983, Table 2.4, total N = 457). Key: MP – Modus Ponens, DA – Denial of the antecedent, AC – Affirmation of the consequent, MT – Modus Tollens.

is endorsed much more often for conditional statements that express causal relationships or are used to make threats or promises (Newstead, Ellis, Evans, & Dennis, 1997). Consider the promise: if you wash my car, I will give you ten dollars. Most people will say that if you suppose the car is not washed then the ten dollars will not be paid (DA). Although not logically implied this makes every sense in terms of the pragmatics of everyday conversation. The speaker wants the car washed and is providing an incentive: it would make no sense to pay someone who did not wash the car. A tip is weaker than a promise pragmatically because the speaker suggests an action will produce a desired outcome but has no actual control over it. An example might be 'if you wash Dad's car, he will give you ten dollars.' The frequency of endorsing even Modus Ponens drops significantly when a tip is substituted for a promise as do *all the other* conditional inferences.

There are many experiments published on how beliefs influence conditional inferences, far too many to discuss here (for reviews, see Evans et al., 1993; Evans & Over, 2004; Nickerson, 2015). All of them show that belief affects conditional reasoning in very significant ways when logically equivalent inferences are presented with different problem content.

7.1.3 The Wason Selection Task

Peter Wason was a British psychologist who is regarded as the father of the modern psychology of reasoning. Most of his influential work was published between about 1960 and 1980, including a book which helped to identify the psychology of reasoning as a research field (Wason & Johnson-Laird, 1972). Of lasting importance has been the invention of several novel tasks for studying reasoning, the most influential of which has been the four-card selection task. Strictly speaking, the selection task does not meet all the definitions of the deduction paradigm as I have given them, as it involves hypothesis testing as well as reasoning. However, the task is focussed on the logic of conditional statements and has been extensively studied by the same research community that has studied conditional inferences and other more conventional reasoning tasks. It has also been used to address broadly the same set of theoretical issues

A typical standard abstract form of the problem is presented in Figure 7.4. The generally accepted cor-

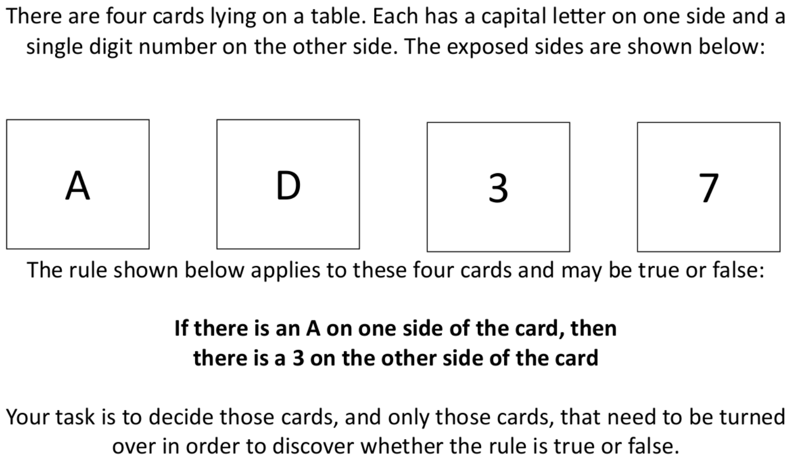

Figure 7.4: The standard abstract Wason selection task.

rect answer is to choose the A and 7 cards, although few participants make these selections. Most people choose either the A card alone, or the A and the 3. Wason pointed out that the conditional statement can only be falsified if there is a card which has an A on one side and does not have a 3 on the other. Clearly, the A card must be turned over because it must have a 3 on the back, and if it does not, disproves the claim. Similarly, the 7 card – which is not a 3 – could have an A on the back, which would also disprove the statement. Turning the 3 is unnecessary, as it cannot disprove the rule. It need not have an A on other side.

Why do people choose A and often 3 and ignore the 7? Wason originally suggested that they had a verification or confirmation bias. They were trying to prove the rule true rather than false, and hence looking to find a confirming combination of A and 3. In support of this account, if you ask people to give written justifications for their choices, they typically say that they were looking for 3 when turning the A and vice versa because this would make the rule true (Wason & Evans, 1975). However, in an early research paper of my own, I showed that this account cannot be right. The trick is to include a negative in the second part of the conditional as in the example shown in Figure 7.5. When the rule says, as in the example, that if there is G on one side of the card, then there cannot be a 4 on the other side of the card, most people choose the G and 4 cards. But these choices do not verify the rule, they correctly falsify it. The combination that falsifies this statement is, of course, G and 4. Once again, participants say that they are turning the G to find a 4 and vice versa (Wason & Evans, 1975) but now in order to make it *false*. It is as though the negative has helped them to understand the logic or the problem.

The effect here is called *matching bias*. People tend to choose the cards which match those named in the conditional, whether or not negations are present. But these negations affect the logic of the task, so it is a puzzling finding. Matching bias is another example of a cognitive bias, like the atmosphere effect, which operates with abstract materials. With the other methods, I showed that the introduction of realistic materials makes a big difference to responding. The same is true of the selection task. It was thought initially that simply using realistic materials made the problem a lot easier with higher rates of correct selections (Wason & Johnson-Laird, 1972). This was known as the *thematic facilitation effect*. However, it was later shown that the versions that make the task really easy include a subtle change to the logic. An example, known as the 'drinking age

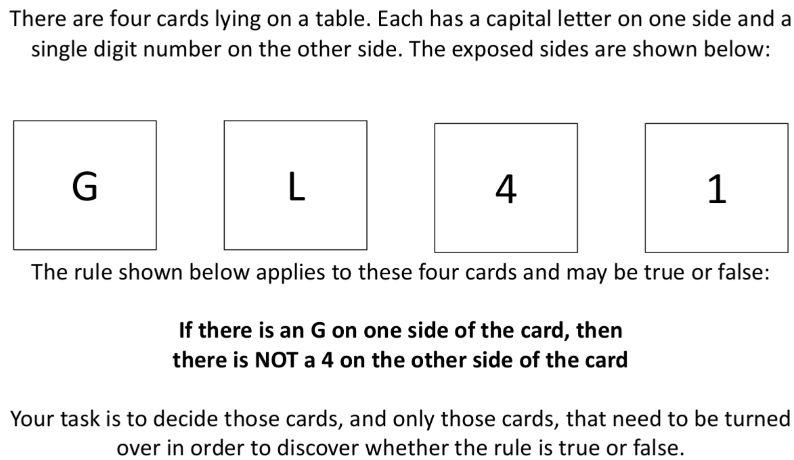

Figure 7.5: The abstract Wason selection task with an added negation.

rule,' is shown in Figure 7.6. People are first of all given a short context. In this case, they are asked to imagine that they are police officers enforcing rules. Then the rule is given which requires beer drinkers to be over 18 years of age. Now most people will check the beer drinkers and those *under 18* of years of age. This is correct as only underage drinkers can violate the rule. Experiments which give the drinking age rule find much higher rates of correct answers than with the standard abstract version.

As later authors pointed out, problems like the drinking age rule change the task from one of indicative logic (concerned with truth and falsity) to one of deontic logic, concerned with obeying rules and regulations. A number of different theoretical accounts have been offered to explain why the deontic version is so much easier. One idea is that we acquire and apply *pragmatic reasoning schemas*: rules which apply in certain contexts and can be instantiated with the content of a particular problem (Cheng & Holyoak, 1985). So people might solve the drinking age rule because they have a permission schema such as 'if an action A is to be taken, then condition C must be filled', which could be instantiated as A = drinking beer and C = 18 years or age or older. The schema tells them violations of this rule occur when the action is taken without the precondition being filled. Other proposals included the use of innate evolved rules for social exchange (Cosmides, 1989), interpreting the problems as a decision-making task in order to maximise perceived benefits (Manktelow & Over, 1991) and the role of pragmatic relevance for different forms of conditional statement (Sperber, Cara, & Girotto, 1995).

7.2 Theoretical Issues in the Psychology of Deduction

Having described the main methods and typical findings in the study of deduction, I now turn to some broader theoretical issue and arguments that have arisen.

7.2.1 How We Reason: Rules or Models?

Despite the frequency of errors and biases, people do show some level of deductive competency on reasoning tasks, especially those of higher cognitive ability (Stanovich, 2011). For example, people endorse far more inferences that are valid than invalid on both syllogistic and conditional inference tasks. Some psychologist have focussed on the competence rather than the errors and asked questions

Imagine you are a police office checking people drinking in a bar. It is your job to ensure that people conform to certain rules. The following cards show on one side what people are drinking and on the other side their age:

Beer	Coke	22 years	16 years

Here is a rule:

If a person is drinking beer then that person must be over 18 years of age

You must decide those cards, and only those cards, that need to be turned over in order to discover whether the rule is being violated.

Figure 7.6: The Deontic selection task: Drinking age rule.

about the mechanisms by which people draw deductions. One approach is often described as **mental logic** but should more accurately be described as mental rule theory (Braine & O'Brien, 1998; Rips, 1994). Traditional logics are usually presented as rules, but other techniques for generating valid deductive inferences are available. The term 'mental logic' was devised to distinguish the logic inside people's heads from that in the philosopher's textbooks. The idea is that ordinary people reason by built-in logical rules. However, the psychological authors were mindful from the start of certain psychological findings. For example, in proposing a mental logic account of conditional inference, psychologists were well aware that people find Modus Ponens a lot easier than Modus Tollens. In standard logic, these would both be primitive rules of equal standing, but that cannot be the case in a mental logic.

Mental logicians have tried to address this problem by proposing only Modus Ponens is a simple rule allowing direct inference. Modus Tollens can be drawn but by an indirect reasoning procedure called *reductio* reasoning. In this kind of reasoning, one makes a supposition p and tries to show that this leads to contradiction, q and not-q. Since a contradiction cannot exist in logic, the supposition must be false, hence not-p follows. Consider our earlier examples of a conditional which applies to cards with a letter on one side and a number on the other

If the letter is B, then the number is 3

If people are told there is a B, then they can immediately apply their built-in rule for Modus Ponens and conclude that there is a 3. When told there is not a 3, however, they do not have a rule for Modus Tollens that can be applied in the same way. Instead, they have to reason as follows:

'If I imagine there is a B on the card, then there must be a 3. But I have been told there is not a 3 which is a contradiction. So I must have been wrong to suppose that there was a B on the card. Hence, I can conclude that there is not a B.'

This indirect reasoning is harder to execute and more prone to errors, explaining the lower acceptance rate of Modus Tollens. Hence a fully specified mental logic account consists of a set of direct rules of inference together with indirect reasoning procedures and can be implemented as a working computer program (Rips, 1994).

For many years now, rule-based mental logics have had a major rival, which is **mental model theory**, championed by Phil Johnson-Laird (1983, 2006; Johnson-Laird & Byrne, 1991) and followed and supported by many other psychologists. This theory also proposes that people are deductively competent in principle, but fallible in practice, but does not rely on the application of mental rules. The core of the theory is the idea that people reason about possibilities, which are states of the world that might be true or false. These possibilities are represented by mental models. When the premises of an argument are presented, people are supposed to construct mental models to represent all possibilities and then reason as follows:

1. If the conclusion is true in all models of the premises, it is necessary
2. If the conclusion is true in some models of the premises, it is possible
3. If the conclusion is true in no models of the premises, it is impossible

Consider the inference of disjunction elimination, discussed earlier. Given the major premise Either the letter is B or the number is 3, people construct the following models as possibilities:

B	
	3
B	3

Each line here represents a separate mental model. So this reads as: one possibility is that there is a B, the second that there is a 3 and the third that there is both a B and a 3.

Now given the minor premise, the number is not 3, the last two models are eliminated leaving only the possibility of B. Hence, people will conclude B as a correct deduction but without having any rule of disjunction elimination for the inference. The model theory of conditionals (Johnson-Laird & Byrne, 2002) is more complex and controversial (Evans, Over, & Handley, 2005). Consider the statement

If there is a B, then there is a 3

The full set of possibilities for this statement, according to the published theory, is the following:

B	3
not-B	3
not-B	not-3

Given these possibilities, the minor premise B eliminates all but the first model, so the conclusion 3 follows (Modus Ponens). The minor premise not-3, eliminates all the last model, so that not-B follows (Modus Tollens). However, like the mental-logic theorists, Johnson-Laird and Byrne were well aware of the relative difficulty of Modus Tollens. So they actually proposed that people initially represent the conditional statement as follows:

[B]	3
...	

The first model means that in all cases of B (meaning of []), there is a 3. The ellipsis '...' means there are other possibilities. So if B is presented, people can immediately do Modus Ponens and conclude 3. However, if not-3 is presented, then they have to 'flesh out' the models to the fully explicit set given above in order to make Modus Tollens. Fleshing out is error prone, so people sometimes fail to make this valid inference. The theory has been applied to many other types of reasoning, including with syllogisms.

I think it fair to say that both mental rule and mental model theory are firmly rooted in the traditional deduction paradigm, as they both put an account of logical competence foremost and deal with effects of beliefs and pragmatics as add-ons. Whether belief rather than logic should be the focus of psychological accounts of reasoning is precisely the issue for researchers in the new paradigm, to which I return later. We next consider **dual-process theory** which is not directly concerned with how logical reasoning occurs, but rather with the idea that such reasoning competes with other kinds of cognitive processes of a more intuitive nature.

7.2.2 Dual-process Theories of Reasoning

In 1982, I published my first book which was a review of the psychology of deductive reasoning. Even at that time, with many of the key studies in the field yet to be conducted, it was clear that logical errors were frequent, and there was much evidence of systematic biases and belief-based reasoning. I was struck by what seemed to be two factors influencing many different reasoning tasks. People's choices were indeed influenced by the logic of the problems, just as had been originally expected, but also by non-logical factors that were completely irrelevant to the task such as atmosphere or matching bias. But two-factor theory is descriptive and provides no real explanation of the cognitive processes that underlie our observations.

An important theoretical leap is from dual sources to dual processes. What if the two factors reflect quite different mental processes? I mentioned earlier that Wason and Evans (1975) observed that when matching bias cued a correct choice on the selection task, people appeared to understand the logic of the problem and the importance of falsification. But in the standard version, they talked as though they could prove the rule true. In this early paper, we suggested a distinction between unconscious Type 1 processes responsible for the matching selections and conscious Type 2 reasoning, which simply served to rationalise or justify the unconsciously cued cards. The radical suggestion was that actual choices were determined by one kind of process but the verbal justifications by something entirely different. The next important step, in work also described earlier, was the belief-bias study of Evans et al. (1983). What we observed there was that syllogistic reasoning was influenced heavily by both the logic of the syllogism and the believability of conclusion. We showed that the two factors were in conflict and that individuals would sometimes go with logic and other times with belief.

The linkage with the Wason and Evans work was not made immediately but in retrospect the view developed was that Type 2, explicit (slow, reflective) reasoning processes were responsible for the preference for valid over invalid conclusions. However, these competed with Type 1 (fast, intuitive) processes, which favoured believable over unbelievable conclusions. So on tasks other than the selection task, at least, Type 2 reasoning could solve problems and not just rationalise intuitions. (Later research showed that Type 2 processes have a role in the selection task choices as well.) As different dual-process accounts were developed over the next quarter of a century, there was a particular emphasis on the idea that Type 2 reasoning was responsible for logically correct answers and Type 1 processing for non-logical effects, such as matching and belief bias (see Evans, 2007; Stanovich, 1999; see also Chapter 10, "Decision Making", for application of dual-process theory to decision making).

A large individual-differences programme conducted by Keith Stanovich and Richard West, who showed that on a wide variety of reasoning and decision-making tasks, cognitive ability or IQ (see Chapter 14, "Intelligence") was strongly correlated with the ability to give the correct answer. The theoretical idea here is that people with higher IQs also have higher working memory capacity and are therefore more able to manipulate mental representations of premises and conclusions in order to reason logically. In fact, the engagement of working memory is now considered a defining feature of Type 2 processing (Evans & Stanovich, 2013). Other methodologies were developed that supported this view. For example, if people are given a very short time to respond, they are less likely to give the logical answer and were more likely to show a bias, such as matching or belief. Similar results occur if a working-memory load has to be carried while reasoning (for reviews see Evans, 2007; Evans & Stanovich, 2013). However, Stanovich and West also showed that general intelligence was not the only individual difference factor in human reasoning. In particular, people vary in **rational thinking disposition** which measures the inclination to accept an intuitive answer or to check it out by high effort reasoning.

Some dual-process theorists have suggested that there are ruled-based (Type 2) and associative (Type 1) processes that operate in parallel (Sloman, 1996) but more popular among deductive-reasoning researchers is the idea that fast intuitive (Type 1) answers come to mind immediately and are subject to checking and possible revision by slower, reflective (Type 2) processes that follow (Evans & Stanovich, 2013; Kahneman, 2011). This leads to the important question of why it is that some intuitive answers are more carefully checked than others. A hot topic in the field right now is whether people's initial answers are accompanied by **feelings of rightness**, which help them decide whether to accept the intuitive answer or whether to check it out with careful reasoning (Hot Topic, see also Chapter 6, "Metacognition").

7.3 The New Paradigm Psychology of Reasoning

The deduction paradigm was developed about 50 years ago to assess the then-prevalent view that logic was the basis of rational reasoning. As evidence of logical errors, cognitive biases, and belief-based reasoning accumulated, this presented a clear problem both for psychologists and for philosophers who became aware of the findings of research on deduction as well as other kinds of human reasoning with similar findings. The problem was that, by the original assumptions, people were turning out to be irrational. Peter Wason, for example, was quite clearly of the view that people were illogical and therefore irrational (see Evans, 2002). In a famous paper, the philosopher Jonathan Cohen argued that people were in fact inherently rational and that psychological experiments could never prove otherwise (Cohen, 1981). He suggested that the experiments were unrepresentative or being misinterpreted. He also pointed out that standard logic, for example, is not the only kind that logicians have offered. There could be alternative normative accounts of how to be rational. A normative theory is one of how people *ought* to reason. Subsequently, a number of psychologists engaged with the issue of what counts as rational reasoning (e.g. Evans & Over, 1996; Stanovich, 1999).

A major issue is whether traditional logic provides the correct standard for human reasoning. One major research programme, that of Mike Oaksford and Nick Chater, has disputed this from the start (Oaksford & Chater, 2001, 2007). Their first important contribution was an alternative normative account of the Wason selection task, arguing that the typical answer can be seen as rational from a decision-making perspective (Oaksford & Chater, 1994). Like Cohen, they took the view that human behaviour must be rationally adapted to the environment and if a standard normative account does not explain it, then we should look for another. They have presented various theories of reasoning tasks based on probability theory and decision theory. Naturally this approach is controversial and has been branded Panglossian by some authors (Stanovich, 1999). Pangloss was a fictional philosopher in a novel by Voltaire who as prone to say 'all is for the best in the best of all possible worlds'!

The essence of the new paradigm is that people naturally reason from their beliefs about the world and that this should not be treated as an error or cognitive bias. Strong deductive reasoning instructions are artificial: they require people to ignore what they believe for the sake of the experiment. Other methods have been explored. For example, people can be asked what inference follows from some information and allowed to express degrees of belief in their conclusions. A key feature of the new paradigm is the proposal of the *suppositional conditional*, also known as the probability conditional (Evans & Over, 2004; Oaksford & Chater, 2001). The conditional statement of standard logic is equivalent to a disjunction. For example,

If the letter is B, then the number is 3

is true except when we have a B and not a 3. Hence it is equivalent in meaning to

Either the letter is not a B or the number is a 3

This is what logicians term the *material conditional*. However, many philosophers have rejected the material conditional as an account of the ordinary conditional of everyday language (Edgington, 1995). This is because it leads to unacceptable inferences. The material conditional, if p then q, is true whenever p is false or q is true. So the following statements must be true

> If President Trump is French, then Paris is the capital of the USA

> If 2+2 = 5, then 3 is a prime number

It is clear that no normal person would endorse these statements.

If the ordinary conditional is not material, then what is it? The philosopher Ramsey famously argued that belief in the ordinary conditional if p then q, is in effect the probability that q will be true if p is (Ramsey, 1931/1990). He also suggested that we do this by adding p to our current stock of beliefs and arguing about q on that basis. This is known as the Ramsey test and we suggested that conditional statements are suppositional – they depend on the supposition of p (Evans & Over, 2004). Let us consider some examples

> If teachers' pay is raised, then recruitment will increase

Many people will agree with this statement or assign a high probability to it. They do this by supposing first that teacher's pay is in fact raised and then using other beliefs to calculate the likelihood that they will prove easier to recruit. They may be aware that recruitment has been difficult in recent years and that one factor is almost certainly that salary levels have fallen behind those of workers in other professions. So, they believe that financial incentive will help address the issue. In doing this, they ignore any beliefs they have about what will happen if pay is *not* increased, which they regard as irrelevant. The Ramsey test is also related to findings with Modus Ponens mentioned earlier. When people believe a conditional statement to be true, they also believe that q is probable when p is assumed and so will readily infer q from p. Consider, however, this statement

> If the global economy grows then there will be less poverty and starvation in the world

The Ramsey test will not produce a high level of confidence in this conditional for many people. They may believe, for example, that the growth in the global economy increases wealth for rich individuals and rich countries but is not likely to be distributed to the third world. where most of the poverty and starvation is concentrated. If their belief in the conditional is low, they will also be reluctant to draw inferences from it, even the apparently obvious Modus Ponens. There is now much evidence that, with real-life conditionals like these, people do indeed assign very similar belief levels to if p then q as they do to the probability of q given p (e.g. Over, Hadjichristidis, Evans, Handley, & Sloman, 2007). People act as though the conditional only applies on the supposition that p is true, in people's minds, and is otherwise irrelevant. This is not consistent at all with the material conditional of standard logic.

The essence of the new paradigm is to view belief-based reasoning as natural and rational, rather than to consider it necessarily a source of bias. The new paradigm is, however, not yet as clearly defined as the old. A lot of authors are pursing alternative normative theories of reasoning to standard logic or seeking explanations in terms of Bayesian decision theory – a system which takes account of subjective beliefs. Others have argued that the new paradigm should concern itself less with normative theory and more with describing what people actually do when they reason. One thing that all agree upon is that the traditional standard logic is neither a good account of how we actually reason, nor of how we ought to reason.

Summary

1. Deduction was studied originally to assess the classical view that logic is the basis for rational human thinking.

2. The traditional deduction paradigm assesses whether people untrained in logic can nevertheless correctly evaluate the validity of deductive inferences. They are instructed to assume some premises are true and to decide if the conclusion necessarily follows.

3. Many psychological experiments were conducted in the second half of the twentieth century within this paradigm, most using one of three methods: syllogistic reasoning, conditional inference or the Wason selection task.

4. Typical findings from these different methods converged. People make many logical errors, shown systematic biases and are strongly influenced by their beliefs about the problem content.

5. People also show a degree of deductive competence. Theories of how this is achieved include the idea that people have built in rules for reasoning – a mental logic – or alternatively that people use mental models to represent logical possibilities.

6. Dual-process theories arose from the observation that the logical deductions people make on these tasks often seem to conflict and compete with cognitive biases. These theories propose the operation of rapid, intuitive Type 1 processes as well as slow, reflective Type 2 processes, the latter engaging working memory.

7. Individual differences studies show that successful reasoning on laboratory tasks is often related to general intelligence or working memory capacity. People also vary in rational thinking dispositions which makes them more or less likely to engage in high effort reasoning. There is also evidence that people experience a feeling of rightness about the intuitive answers that come easily to mind and that they are less likely check answers by reasoning if this feeling is strong.

8. The new paradigm psychology of reasoning has arisen in the past twenty years or so as many psychologists did not accept that illogicality implied irrationality. According to this view, the frequent intrusion of beliefs into laboratory reasoning tasks where they are defined as irrelevant tells us something important about human reasoning. The paradigm assumes that belief-based reasoning is both normal and adaptive in everyday life.

Review Questions

1. Are people rational in their reasoning, given their apparent illogicality in laboratory tests of deductive reasoning?

2. Is it more plausible that people make deductions by accessing logical rules and performing mental proofs, or by simply reasoning about what is or is not possible given the premises of an argument?

3. Are cognitive biases observed in the laboratory a real concern for real world reasoning and decision making? For example, is it likely that belief bias prevents us from fairly assessing evidence for rival hypotheses?
4. Are dual-process accounts inherently plausible? That is, can you think of many examples in everyday life where intuitions might come into conflict with reflective reasoning?
5. Is it safe and appropriate to rely on a strong feeling that you are drawing the correct inference or making a good decision?

Hot Topic: Feelings of rightness in reasoning and implications for dual-process theory

Jonathan Evans

The standard dual-process approach assumes that an intuitive answer to reasoning (and decision making) problems comes to mind quickly and rapidly due to Type 1 processing. It is then subject to checking by Type 2 processes, which may rationalise the intuitive answer or substitute more complex reasoning to provide a different answer. An important question is why we sometimes rely on the initial intuition with minimal reasoning and other times engage Type 2 processes. Some known factors are the processing style of the individual and time available for thinking. However, an additional factor has been proposed by Valerie Thompson and her colleagues – metacognitive feelings. The initial intuition comes with a *feeling of rightness (FOR)*, which could determine whether we accept it or expend effort on reasoning (Thompson, Prowse Turner, & Pennycook, 2011). Thompson has claimed in several papers that FOR is the key factor in determining whether extensive Type 2 processing occurs.

Thompson invented a methodology called the two-response task (Thompson et al., 2011). Participants are given a reasoning or decision task and asked to generate an initial answer as quickly as possible without reflecting on it. They then rate the degree to which they are confident that the answer is correct – the FOR. Following this, they are then asked to think again about the problem, taking as long as they like. After this, they again give a response to the problem which may or may not be the same as first answer. Using a range of different tasks, the following pattern was established: when FOR is high, the initial answer tends to be given quickly, rethinking time tends to be short and second response usually matches the first. In other words, when we are intuitively convinced of our original answer, we expend little effort in trying to check or correct it. Conversely, when FOR is low, we are more likely to take time rethinking the problem and to change our original answer.

There are some unresolved difficulties with this account. First, none of the studies on this report a relation between FOR and actual accuracy. We are no more likely to be confident of a correct answer than a biased one and we are just as likely to change a right answer to a wrong one after reflection as the other way around. So if FOR has evolved to help us make good decisions, using our cognitive resources effectively, why is it not helping us identify errors? In fact, the opposite seems to be true. Matching bias and belief bias, for example, have been shown to be supported by false feelings of rightness. Matching cards and believable conclusions *feel* right, even when they are wrong. Another difficulty is that there is no direct evidence for a casual connection between FOR and Type 2 reasoning as Thompson and colleagues claim. Everything is in fact correlational. All we really know is that answers with high FOR are also made more quickly, thought about less and changed less often.

There are also some recent empirical findings which raise difficulties for the dual-process story here. On relatively simple tasks, when a correct choice is put in conflict with a bias, there is evidence that this conflict is detected very rapidly by the brain, indicating that some kind of 'logical intuition' (De Neys, 2014) is available to conflict with the bias. Recently, using the two-response task with syllogistic reasoning, Bago and De Neys (2017) showed that most people who are correct at time 2 were also correct at time 1. There was little evidence for people correcting intuitive responses by a period of reflection, as might be expected with Type 2 intervention. However, it is possible that studies of this kind provide a misleading impression, as the tasks are relatively simple. Hence, it could be that 'logical intuitions' arise from Type 1 rather than Type 2 processing as these solutions do not require the engagement of working memory on the tasks used (Evans, 2018). As is the nature of hot topics, there is as yet no clear resolutions to the questions I have raised, and research continues.

References

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. doi:10.1016/j.cognition.2016.10.014

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*(2), 169–187. doi:10.1080/13546783.2013.854725

Evans, J. St. B. T. (2018). Dual-process theory: perspectives and problems. In W. De Neys (Ed.), *Dual process theory 2.0* (pp. 137–155). London: Routledge. doi:10.4324/9781315204550-1

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. doi:10.1016/j.cogpsych.2011.06.001

References

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. doi:10.1016/j.cognition.2016.10.014

Braine, M. D. S., & O'Brien, D. P. (1998). *(Eds) Mental logic*. Mahwah, New Jersey: Lawrence Erlbaum Associates. doi:10.4324/9781410603005

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

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*(3), 317–370. doi:10.1017/s0140525x00009092

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? *Cognition*, *31*(3), 187–276. doi:10.1016/0010-0277(89)90023-1

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*(2), 169–187. doi:10.1080/13546783.2013.854725

Edgington, D. (1995). On conditionals. *Mind*, *104*(414), 235–329. doi:10.1093/mind/104.414.235

Evans, J. St. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, *128*(6), 978–996. doi:10.1037//0033-2909.128.6.978

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Hove, UK: Psychology Press. doi:10.4324/9780203947487

Evans, J. St. B. T. (2018). Dual-process theory: perspectives and problems. In W. De Neys (Ed.), *Dual process theory 2.0* (pp. 137–155). London: Routledge. doi:10.4324/9781315204550-1

Evans, J. St. B. T. , Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*(3), 295–306. doi:10.3758/bf03196976

Evans, J. St. B. T. , Handley, S. J., Harper, C., & Johnson-Laird, P. N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(6), 1495–1513. doi:10.1037//0278-7393.25.6.1495

Evans, J. St. B. T. , Newstead, S. E., & Byrne, R. M. J. (1993). *Human Reasoning: The Psychology of Deduction*. Hove, UK: Erlbaum.

Evans, J. St. B. T. , & Over, D. E. (1996). *Rationality and Reasoning*. Hove: Psychology Press. doi:10.4324/9780203027677

Evans, J. St. B. T. , & Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Evans, J. St. B. T. , Over, D. E., & Handley, S. J. (2005). Supposition, extensionality and conditionals: A critique of Johnson-Laird & Byrne (2002). *Psychological Review*, *112*(4), 1040–1052. doi:10.1037/0033-295x.112.4.1040

Evans, J. St. B. T. , & Stanovich, K. E. (2013). Dual process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. doi:10.1177/1745691612460685

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, *69*(4), 366–378. doi:10.1037/h0042043

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

Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199551330.001.0001

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: a theory of meaning, pragmatics and inference. *Psychological Review*, *109*(4), 646–678. doi:10.1037//0033-295x.109.4.646

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux. doi:10.4324/9781912453207

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*(4), 852–884. doi:10.1037//0033-295x.107.4.852

Manktelow, K. I. (1999). *Reasoning and Thinking*. Hove, UK: Psychology Press. doi: 10.4324/9780203015568

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, *39*(2), 85–105. doi:10.1016/0010-0277(91)90039-7

Newstead, S. E., Ellis, C., Evans, J. St. B. T. , & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking & Reasoning*, *3*(1), 49–76. doi:10.1080/135467897394428

Nickerson, R. B. (2015). *Conditional reasoning*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780190202996.001.0001

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631. doi:10.1037//0033-295x.101.4.608

Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*, *5*(8), 349–357. doi:10.1016/s1364-6613(00)01699-5

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198524496.001.0001

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T. , Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*(1), 62–97. doi:10.1016/j.cogpsych.2006.05.002

Ramsey, F. P. (1990). General propositions and causality (original publication, 1931). In D. H. Mellor (Ed.), *Philosophical papers* (pp. 145–163). Cambridge: Cambridge University Press.

Rips, L. J. (1994). *The Psychology of Proof*. Cambridge, MA: MIT Press.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22. doi:10.1037//0033-2909.119.1.3

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

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahway, NJ: Lawrence Elrbaum Associates. doi:10.4324/9781410603432

Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195341140.001.0001

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. doi:10.1016/j.cogpsych.2011.06.001

Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*(2), 141–154. doi:10.1016/0010-0277(74)90017-1

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content*. London: Batsford.

Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology*, *16*, No 102.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology*, *18*(4), 451–460. doi:10.1037/h0060520

Glossary

belief bias A disposition to accept an argument as valid because you believe the conclusion. 116

deduction paradigm A method and tradition that involves asking people untrained in logic to evaluate the validity of arguments. 114

dual-process theory The claim that there are two distinct types of cognitive processing: Type 1 (rapid, intuitive) and Type 2 (slow and reflective). 123

feeling of rightness (FOR) The feeling of confidence that someone has in a quick intuitive answer when reasoning or making judgements. 125

mental logic The theory that there is a logic in the mind for reasoning comprised of inference rules and mechanisms for applying them. 122

mental model theory The theory that we reason about logical possibilities, represented by mental models, without need for rules of inference. 122

new paradigm (psychology of reasoning) A method and theory which allows for beliefs to exert a rational influence on human deductive reasoning. 114

rational thinking disposition An individual difference measure related to personality or cognitive style. It measures the extent to which people rely on intuitions or check them out by reasoning. 124