NeuroImage 275 (2023) 120180

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Probabilistic and deductive reasoning in the human brain

Lupita Estefania Gazzo Castañeda*, Benjamin Sklarek, Dennis E. Dal Mas, Markus Knauff

Experimental Psychology and Cognitive Science, Justus Liebig University Giessen, Otto-Behaghel-Str. 10F, Giessen 35394, Germany

ARTICLE INFO

Keywords:
Reasoning
Conditionals
Deductive reasoning
Probabilistic reasoning

ABSTRACT

Reasoning is a process of inference from given premises to new conclusions. Deductive reasoning is truth-preserving and conclusions can only be either true or false. Probabilistic reasoning is based on degrees of belief and conclusions can be more or less likely. While deductive reasoning requires people to focus on the logical structure of the inference and ignore its content, probabilistic reasoning requires the retrieval of prior knowledge from memory. Recently, however, some researchers have denied that deductive reasoning is a faculty of the human mind. What looks like deductive inference might actually also be probabilistic inference, only with extreme probabilities. We tested this assumption in an fMRI experiment with two groups of participants: one group was instructed to reason deductively, the other received probabilistic instructions. They could freely choose between a binary and a graded response to each problem. The conditional probability and the logical validity of the inferences were systematically varied. Results show that prior knowledge was only used in the probabilistic reasoning group. These participants gave graded responses more often than those in the deductive reasoning group and their reasoning was accompanied by activations in the hippocampus. Participants in the deductive group mostly gave binary responses and their reasoning was accompanied by activations in the anterior cingulate cortex, inferior frontal cortex, and parietal regions. These findings show that (1) deductive and probabilistic reasoning rely on different neurocognitive processes, (2) people can suppress their prior knowledge to reason deductively, and (3) not all inferences can be reduced to probabilistic reasoning.

1. Background and motivation

For decades, cognitive psychologists studied human deductive reasoning by asking people to draw logically valid conclusions from given premises (e.g., Johnson-Laird & Byrne, 1991; Rips; 1994; Braine, 1978). An inference was considered rational if it conformed with the norms of classical logic, otherwise irrational. The roots of classical logic can be traced back to Aristotle, and latter to the pioneering work by the eminent mathematicians and logicians George Boole (1854) and Gottlob Frege (1879). The first psychological experiments on deductive reasoning were conducted by Störring (1908) and then by Wilkins (1928) and Woodworth and Sells (1935). In the following, we refer to this research as the *deductive paradigm*. The results of experiments in this paradigm show that humans make many deductive reasoning errors, they frequently draw inferences that are deductively invalid, and they are often biased by factors that are irrelevant according to the principles of classical logic (Evans, Newstead, & Byrne, 1993).

Today, many reasoning researchers argue that the deductive paradigm is inadequate for studying human thinking. Classical logic is not the right standard for human rationality, they argue, and hence the earlier results do not prove people to be irrational. They just show that the wrong normative standards were used. As an alternative normative framework these researchers suggest probability theory, the calculus of uncertain reasoning, instead of logic, the calculus of certain reasoning (Elqayam, Bonnefon, & Over, 2016; Oaksford & Chater, 2020; Over, 2009; Evans, 2012). The foundations of this normative framework go back to the groundbreaking work by Reverend Thomas Bayes (1763) and Frank P. Ramsey (1990). In a typical probabilistic reasoning experiment, people are asked to use their prior knowledge to evaluate how probable a conclusion is given what they know about the content of the premises. They do not just choose between "true" or "false", as in the deductive paradigm, but rather give a probability rating between 0 and 100%, from highly improbable to highly probable. In the following, we refer to this research as the *probabilistic paradigm*.

The probabilistic paradigm has had a huge impact on cognitive psychology in the last decades. More and more studies have begun to investigate uncertain reasoning, and even theories initially developed for deduction have proposed accounts to explain uncertain reasoning and the consideration of probabilities (e.g., Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999; Johnson-Laird, Khemlani, & Goodwin, 2015). The growing interest in probabilities, however, also has the consequence that some researchers completely deny deductive reasoning as a faculty of the human mind. Oaksford (2015) even argues that what looks like

---

* Corresponding author.
E-mail address: Estefania.Gazzo@psychol.uni-giessen.de (L.E. Gazzo Castañeda).

https://doi.org/10.1016/j.neuroimage.2023.120180.
Received 13 April 2023; Received in revised form 17 May 2023; Accepted 18 May 2023

deductive inferences are actually probabilistic inferences, only with extreme probabilities. But is all human reasoning indeed probabilistic? Or do people also reason deductively? The goal of the present paper is to answer these questions by measuring the cortical activity of participants while they solve reasoning problems either with a deductive or with a probabilistic instruction. The results disagree with the assumption of the probabilistic approach to reasoning: not all reasoning processes are probabilistic, and not all deductive reasoning processes can be reduced to probabilistic reasoning. People can and actually do reason both deductively and probabilistically.

2. Deductive reasoning

Most cognitive experiments on deduction, that is truth-preserving reasoning, use binary extensional logic as the normative standard. Participants are usually confronted with conditional inference tasks that consist of two premises and a conclusion they have to evaluate. Participants are instructed to assume the premises are true and to indicate what necessarily follows from these premises. An example (cf. De Neys, Schaeken, & d'Ydewalle, 2002):

If a person does sports (*p*), then the person loses weight (*q*).
A person does sports (*p*).
―――――――――――――――
The person loses weight (*q*).

The sentences above the line are the premises, the statement below the line is the conclusion to-be-verified. The first premise is also called the major premise and the second one the minor premise. The major premise contains the conditional rule consisting of an if-part (antecedent) and a then-part (consequent). To provide their answers, participants can choose between two responses: they can say that the conclusion is true or that it is false, nothing else. This dichotomous response format results from the binary logic that underlies those experiments: in classical logic a conclusion can only be true or false and nothing in between. In the example the conclusion is true because

If *p*, then *q*
p
Therefore, *q*

is a logically valid inference called modus ponens (MP). Another valid inference is modus tollens (MT):

If *p* then *q*
Not-*q*
Therefore, not-*p*.

Invalid inferences are the affirmation of the consequent (AC):

If *p* then *q*
q
Therefore, *p*

and the denial of the antecedent (DA):

If *p* then *q*
Not-*p*
Therefore, not-*q*.

The validity of an inference depends only on the logical structure of the inference. The concrete content of the conditional is irrelevant. Dozens of experiments have shown that the MP inference is relatively easy for most people, while the MT inference is harder, and that people often accept AC and DA inferences (Evans et al., 1993), which are mistakes in classical logic.

Reasoning researchers have suggested different cognitive theories to explain this pattern of results in human deductive reasoning (Johnson-Laird & Byrne, 1991; Evans et al., 1993; Knauff, 2006; Knauff and Spohn, 2021). Mental rules theories adopt a *syntactic* point of view (e.g., Rips, 1994; Braine, 1978; Braine & O'Brien, 1991, 1998; Braine, Reiser, & Rumain, 1984). The main assumption is that people reason deductively by applying abstract reasoning rules to the premises of reasoning problems. Such mental rules are similar to those of formal classical logic. People consider the to-be-verified conclusion of an inference task to be true if there is a match between it and the conclusion reached by the application of the mental rules. Due to its language-like nature, mental rules theories have often been called linguistic or syntactic approaches to reasoning (see Goel & Dolan, 2001; Goel, Buchel, Frith, & Dolan, 2000). Accordingly, some support for this approach comes from neuroimaging studies in the deductive paradigm that reported activations in areas related to language, such as the left inferior gyrus, the lingual gyri, the left middle/superior temporal lobe, and the left inferior temporal lobe (e.g., Goel et al., 2000; Goel & Dolan, 2003; Luo, Tang, Zhang, & Stupple, 2014; Prado, Chadha, & Booth, 2011; Reverberi et al., 2010). However, not all studies have found activation in these areas (e.g., Noveck, Goel, & Smith, 2004; Knauff, Mulack, Kassubek, Salih, & Greenlee, 2002). One possible explanation for these inconsistent findings is that these language-related activations are not reasoning specific, but only the result of the comprehension and encoding of verbal information into mental representations (e.g., Coetzee & Monti, 2018; Monti & Osherson, 2012).

Another influential theory is the mental model theory of reasoning (Johnson-Laird & Byrne, 1991, 2002). The key idea is not to use formal inference rules but the meaning and interpretation of the premises, which are used to construct mental models (Knauff, 2006; Johnson-Laird, 2001; Yang & Johnson-Laird, 2000). These mental models represent what the premises describe. People can mentally inspect and vary these models to arrive at a valid conclusion.[1] An important characteristic of mental models is that they are spatial representations which are more abstract than visual images and more concrete than language-based representations (Gazzo Castañeda & Knauff, 2013; Knauff & Johnson-Laird, 2002; Knauff & May, 2006). The cognitive processes responsible for deductive reasoning should therefore resemble those of spatial reasoning (Knauff, 2013). The account is supported by several studies in the deductive paradigm that found activations in parietal areas that are related to spatial reasoning, such as superior and inferior parietal regions and the precuneus (e.g., Knauff et al., 2002; Mackey, Singley, & Bungle, 2013; Rodriguez-Moreno & Hirsch, 2009). These parietal areas are consistently activated during relational reasoning (e.g., Knauff, 2009; Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003; Fangmeier, Knauff, Ruff, & Sloutsky, 2006; Goel & Dolan, 2001; Wertheim & Ragni, 2018), while the results for conditional reasoning are heterogeneous. Some studies found activations in spatial areas (Noveck et al., 2004), but others not (e.g., Prado, Van der Henst, & Noveck, 2010; Reverberi et al., 2010). In a meta-analysis, Prado et al. (2011) found consistent activations for conditional reasoning in the left medial frontal gyrus, the left precentral gyrus, and the posterior parietal cortex (PPC). In particular the PPC could be linked to spatial reasoning (cf. Knauff, 2013). However, the PPC activation was left lateralized and centered around the angular gyrus, which according to Prado et al. (2011) is related to verbal processing and not spatial reasoning. A more recent meta-analysis by Wertheim and Ragni (2020) again found parietal activations, but only for conditionals with abstract content (e.g., "If there is a box, then there is a ball"). Conditionals with meaningful content (e.g., "If Ann eats too much sugar, then she will gain weight"), however, led mostly to activations in the left frontal lobes, which were also active during reasoning with abstract conditionals.

---

[1] Mental models theory was initially created as a theory of deduction, but has been adapted in the last years to also cover probabilistic and everyday reasoning. In this paper, however, we focus only on the initial theory of deduction.

3. Probabilistic reasoning

The starting point of probabilistic reasoning is that prior knowledge is essential for human reasoning (Oaksford & Chater, 2020). When we reason in our daily life, we seldom assume the information we get is either true or false, but rather assign degrees of belief to the content of an inference. Therefore, most researchers in this field explore how people reason with uncertainties and use Bayesian probability theory as the normative standard (e.g., Oaksford & Chater, 2007; Evans & Over, 2004; Pfeifer, 2013). Take, for instance, our sports example, in which the conclusion that a person loses weight after doing sports is deductively true. It is a valid inference, but when people are asked to evaluate this conclusion, they often hesitate to accept it (e.g., Byrne, 1989; Cummins, 1995; De Neys, Schaeken, & d'Ydewalle, 2003; Gazzo Castañeda & Knauff, 2018, 2021). Probabilistic theories explain this by arguing that people use their prior knowledge and know, for instance, that although sport usually leads to weight loss, this is not always the case. They know that *disabling conditions*, such as wrong exercises or bad nutrition, can prevent people from losing weight even though they do sports (see Cummins, Lubart, Alksnis, & Rist, 1991; Cummins, 1995; De Neys et al., 2003). To capture this uncertainty of everyday reasoning, in probabilistic theories conditionals are understood as the *conditional probability* of q given p (Evans, 2012; Evans, Handley, & Over, 2003; Oaksford & Chater, 2020). This is described by the Ramsey test (1929/1990). Participants first assume that p holds (here: that a person does sports). Based on this assumption and their prior knowledge, they then calculate how probable q is (here: losing weight). The more disabling conditions they know, the lower the conditional probability. And the lower the conditional probability, the less probable it is that MP and MT conclusions are accepted (Evans & Over, 2004; Weidenfeld, Oberauer, & Hörnig; 2005; but see Oaksford, Chater, & Larkin, 2000). Therefore, to be able to give uncertain responses, in the probabilistic paradigm participants usually have to provide their inferences on a scale (often ranging from 0% to 100%) and indicate how strongly they accept a conclusion. This differs strongly from experiments on deduction, where participants typically can only choose between two response options: true and false.

Since prior knowledge is essential for probabilistic reasoning, it is reasonable that the underlying processes of probabilistic reasoning are related to retrieval from memory. In neuroimaging studies this should result in activations in hippocampal areas (e.g., Greicius et al., 2003; Lepage, Habib, & Tulving, 1998). However, until now there are no neuroimaging studies that explicitly address probabilistic reasoning. There is, on the one hand, a study by Schwartz, Epinat-Duclos, Léone, and Prado (2017), in which children were asked to evaluate conclusions either according to their logical validity or according to their likelihood. The question about the likelihood could be considered as a kind of probabilistic reasoning, which, interestingly, led to activations in right frontal and bilateral parietal regions. However, participants were always instructed to consider the premises as true – also in the likelihood condition – which is a quite logical and very "un-probabilistic" instruction. On the other hand, there is indirect evidence from neuroimaging studies on *inductive reasoning* that dealt with some kind of "probabilistic" reasoning (e.g., Goel, Gold, Kapur, & Houle, 1997; Goel & Dolan, 2004; Osherson et al., 1998; Parsons & Osherson, 2001). In these studies, activation was found in regions in the left medial frontal gyrus (Goel et al., 1997) and the left dorsolateral frontal gyrus (Goel & Dolan, 2004; Osherson et al., 1998), but also in areas related to memory retrieval, such as the posterior cingulate, parahippocampal regions, and medial temporal areas (Parsons & Osherson, 2001). From a probabilistic point of view, however, these neuroimaging studies have several disadvantages. Participants did not respond on a scale, which is essential for mapping probabilities, but rather gave binary responses (e.g., plausible vs. implausible). And, most importantly, the inductive inferences were very different from the probabilistic reasoning problems currently discussed in the probabilistic paradigm. They did *not* measure the neural correlates of people using their prior knowledge to reason probabilistically with conditional inferences that can be valid or invalid from a deductive point of view. Instead, they focused on other kinds of probabilistic reasoning, such as generalizations from specific observations to general rules (e.g. "George was a woolly mammoth. George ate pine cones. All woolly mammoths ate pine cones" in Goel et al., 1997; see also Goel & Dolan, 2004 and Kochari, van Rooij, & Schulz, 2020). Overall, then, there is still no single brain imaging study that uses the paradigms currently routinely employed in behavioral studies on probabilistic reasoning.

4. Deductive vs. probabilistic reasoning

A main difference between the deductive and probabilistic paradigms is thus the role of prior knowledge. While prior knowledge is essential for probabilistic reasoning, it is (by definition) irrelevant for deduction. In fact, for probabilistic reasoning, prior knowledge is necessary to calculate the probabilities of premises and hence evaluate conclusions. In contrast, in deductive reasoning the validity of a conclusion only depends on the structure of the premises. Hence, for deciding whether a conclusion is deductively valid, people have to ignore what they know about the content and focus solely on the logical structure of the inference.

Researchers from the probabilistic camp deny that this kind of "logical" reasoning is possible. People cannot suppress what they know, it always affects their inferences. Therefore, according to Oaksford, one of the main advocates of this approach, every kind of reasoning is probabilistic and always involves prior knowledge – even when people reason with abstract problems, i.e., problems with a content about which they do not have any particular prior knowledge (see Oaksford, 2015; see also Oaksford & Chater, 2010). People can achieve "more logic like processing" by inhibiting prior knowledge (Oaksford & Chater, 2012, p. 21), but this reasoning is still probabilistic, just now with probabilities of 0 or 1 (Oaksford, 2015, p. 4). Based on this assumption, Oaksford (2015) argues that the existing studies on deduction actually did not measure deductive reasoning, but just a special form of probabilistic reasoning. In other words, there is no such thing as deductive reasoning, but only probabilistic reasoning that can lead to responses that look like the results of deductive reasoning. Some neuroimaging studies seem to support this view, by showing that deductive reasoning is often accompanied by activations in areas related to inhibitory control (e.g., Goel et al., 2000; Houdé et al., 2000, 2001; Houdé & Borst, 2015; Prado & Noveck, 2007; see also De Neys, Vartanian, & Goel, 2008). However, there is one difficulty with this interpretation. Even if deductive reasoning is accompanied by the inhibition of prior knowledge, this does not imply that the reasoning process is probabilistic. Deductive reasoning can still rely on specific cognitive processes that differ from those of probabilistic reasoning.

The goal of the present research is to uncover the neural correlates of deductive and probabilistic reasoning in the human brain. Our hypothesis is that people do not reason only probabilistically, but that they can actually do both deductive and probabilistic reasoning. Yes, we agree that prior knowledge and memory retrieval are essential in everyday reasoning. We live in an uncertain world where the information we get is only true to a certain degree. Hence, when people have to reason deductively, they may indeed have to inhibit their prior knowledge. However, once prior knowledge has been inhibited, deductive reasoning can still rely on specific cognitive processes that differ from those of probabilistic reasoning. People can thus reason one or the other way, depending on the given information, the context, the demands of the situation, the availability and suitability of prior knowledge, etc. If a task allows considering prior knowledge, people may indeed retrieve knowledge about the contents of inferences from memory to decide how strongly they accept a conclusion. If, however, a task demands to ignore content and to consider only the structure of inferences, then people can concentrate on the logical form of an argument to reason deductively.

To test this view on human reasoning, we conducted an fMRI experiment in which

1. two groups of participants were instructed either to reason deductively or to reason probabilistically. This variation should allow us to directly compare the neural correlates of the two kinds of reasoning;
2. all participants could freely choose between two response formats. They could either give a binary answer (“yes” or “no”) on a dichotomous response format or a graded response on a scale. This factor should help us to explore whether people indeed always consider uncertainties during reasoning, as suggested by probabilistic theories;
3. the content of the inferences was systematically varied. Two thirds of the problems had a content that people definitely had prior knowledge about. In half of these problems the conditional had a high conditional probability and in the other half a low conditional probability. The last third of the problems was abstract, i.e., people did not have any prior knowledge on the conditional relationship and thus no probability information they could use for the inference. This factor should help us to explore whether people still prefer to give a graded response, if possible, as suggested by the proponents of the probabilistic theories.

For all three variations, we had clearly testable predictions, both on the behavioral and the cortical level. On the behavioral level, we expected that participants under deductive instructions should be able to focus on the structure of the inference and inhibit their prior knowledge about the content of the inference. As a result, participants should prefer giving binary responses under deductive instructions, whereas they should prefer graded responses under probabilistic instructions. We also hypothesized that the contents of the reasoning problems affect inferences under probabilistic instructions, but not under deductive instructions.

On the cortical level, we expected different activations for deductive and probabilistic reasoning. In deductive reasoning, people have to suppress prior knowledge, which should be accompanied by increased activity in brain regions related to inhibition. These are the right lateral prefrontal cortex (Goel et al., 2000; Goel & Dolan, 2003; see also Houdé et al., 2000, 2001; Houdé & Borst, 2015), mid-dorsolateral prefrontal cortex (Prado & Noveck, 2007; see also Knoch, Pascual-Leone, Meye, Treyer, & Fehr, 2006), the right inferior frontal cortex (IFC; e.g., Tsujii, Masuda, Akiyama, & Watanabe, 2010; Aron et al., 2003), but also the anterior cingulate cortex (ACC), which has been related to cognitive control and conflict monitoring (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004). We should also find neural activity in areas that were already identified in previous studies on deductive reasoning (cf. Prado et al., 2011). Since the previous research is still equivocal, we predict additional activity either in language-related or in spatial brain areas or in both. In any case, this activity should only be measurable during deductive but not during probabilistic reasoning.

In probabilistic reasoning, in contrast, our most important prediction is that people will consider their prior knowledge, which should be accompanied by neural activity in areas of memory retrieval such as the hippocampal regions (e.g., Addis, Wong, & Schacter, 2007; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Svoboda, McKinnon, & Levine, 2006). Crucially, these brain areas should be only active during probabilistic but not during deductive reasoning. Conversely, the “inhibition areas” should be only active during deductive but not during probabilistic reasoning. We now report the experiment that tested these hypotheses. The pattern of results is complex but clearly supports our hybrid view on the cognitive and neural basis of human reasoning.

5. Methods

5.1. Participants

Forty-four healthy right-handed adult participants took part in the study. Four participants were excluded from the analysis. One because of dizziness, one deliberately ignored the instructions, and two participants reported a malfunction of the response device. The final sample thus consisted of 40 participants (23 female, 17 male). Participants were on average 24.05 years old ($SD = 2.68$), had normal or corrected-to-normal vision, and gave fully informed consent. Participants were randomly assigned to two groups: twenty participants were instructed to reason deductively, the other twenty received probabilistic instructions. In the following, we refer to the two groups of participants as the *deductive reasoning group* and the *probabilistic reasoning group* (see Materials and Design section). All participants received a monetary compensation for taking part in the experiment. Written consent was obtained from each participant. The study was approved by the Local Ethical Committee of the Department of Psychology and Sports Science at the University of Giessen.

5.2. Materials and design

All participants had to solve 96 conditional inference problems and 24 baseline problems. The inference problems consisted of forty-eight valid and forty-eight invalid inferences. The valid inferences were MP and MT inferences. The invalid inferences had the same structure as the MP and MT inferences, but with the opposite conclusion (“If *p*, then *q; p*; therefore not-*q*” and “If *p*, then *q*; not-*q*; therefore *p*”). The inference problems consisted thus of a major premise containing the conditional rule, a minor premise referring either to the antecedent or the consequent of the conditional, and a conclusion. The conditional probability was varied in the first premise. Eight conditional premises had a *high conditional probability*, eight had a *low conditional probability*, and eight had an *abstract content*. The conditionals with high and low conditional probabilities were taken from De Neys et al. (2002) and Verschueren et al. (2005). Verschueren et al. directly asked participants for the likelihood of *q* given *p*, while De Neys et al. asked their participants to generate disabling conditions. We therefore used the number of disabling conditions to classify the conditionals of the latter study as having high or low conditional probability. Here are examples for the three types of conditionals:

> High conditional probability: “If the person jumps into the swimming pool, then the person gets wet.”
>
> Low conditional probability: “If the person sits in the draught, then the person catches a cold.”
>
> Abstract: “If the box is empty, then the box has stars on it.”

The conditionals with high and low conditional probabilities thus described relationships for which participants had prior knowledge, while abstract problems described relationships for which they did not have prior knowledge. For the low conditional probabilities, participants can usually generate more disabling conditions (e.g., that the person wears warm clothes, has a good immune system, is used to cold weather, took vitamins for precaution, the draught is not so strong, etc.) than for the high probability conditionals (there are only very few exceptions to getting wet when jumping into a swimming pool). All conditionals (low, high, and abstract) were presented four times: in a valid and invalid MP inference as well as in a valid and invalid MT inference. Participants in the deductive reasoning group were instructed to indicate “*whether the conclusion follows logically from the previous premises*” and instructed that a conclusion is *“logically valid if it follows necessarily from the premises*”.

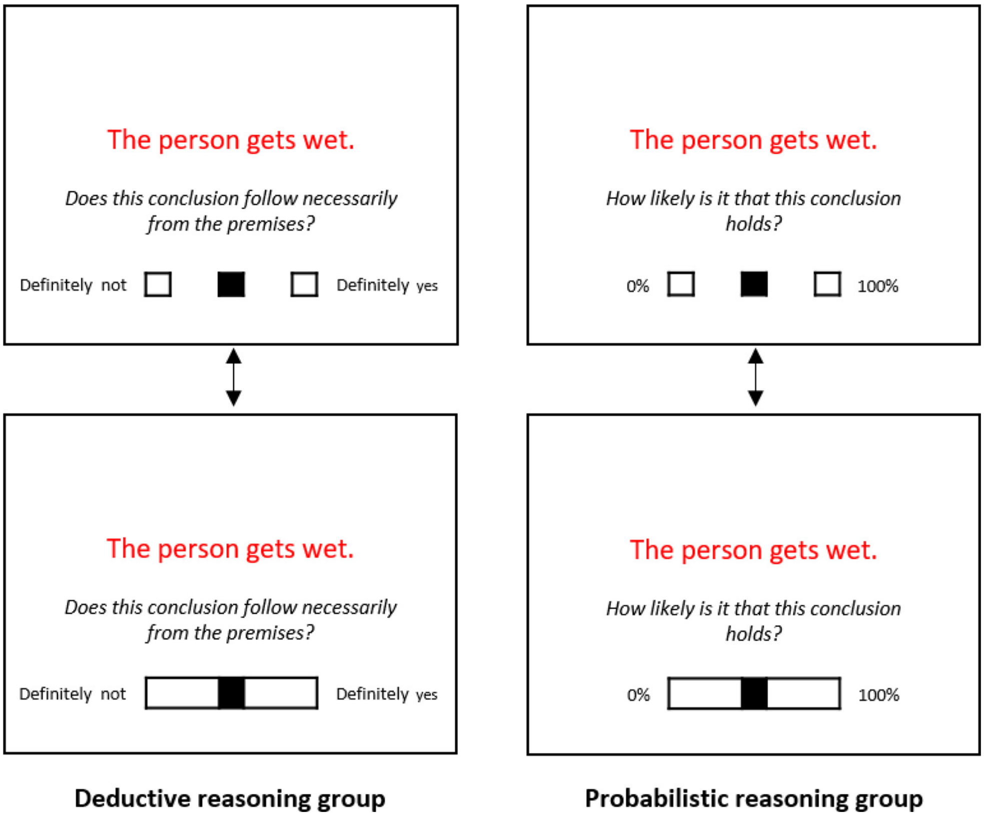

Fig. 1. Flexible response format. Participants could expand the dichotomous response format into a scale and then switch from one to the other format as many times as they wanted. The figure only shows the screen during conclusion evaluation. The entire stimuli presentation including premises is presented in Fig. 2.

Therefore, they should "*consider only the information in the premises*". Participants in the probabilistic reasoning group, in contrast, had to indicate "*how likely it is that the conclusion holds*". They should do so "*as in everyday life and spontaneously*". Therefore, they were told that "*they could consider their own experiences and beliefs*".

All problems were presented with the flexible response format that is illustrated in Fig. 1. Initially, participants saw a dichotomous response format, with the labels "*definitely not*" and "*definitely yes*", in the deductive reasoning group, and the labels "*0%*" and "*100%*", in the probabilistic reasoning group. But then, if wanted, they could press a key in order to expand this dichotomous format into a 7-point Likert scale, with the same labels at the poles. They could switch between the two formats as often as they wanted and give their answer by moving a black square to the right or left, until the desired position. We decided to always present the dichotomous format first, to see whether people indeed feel the urge to give graded responses, as suggested by probabilistic theories.

For the baseline task we created a simple reading and comparison task. First, we presented to the participants two sentences whose content we had previously used in the inference task (e.g., "The person jumps into the swimming pool", "The person gets wet"). Then, we repeated the first of these two sentences (e.g., "The person jumps into the swimming pool"). And finally, we presented either the second of these two sentences (e.g., "The person gets wet") or its negation (e.g., "The person does not get wet"). Participants had to indicate whether or not (deduction group) or to what extent (probabilistic group) the first two sentences were repeated correctly. The baseline problems were also presented with the flexible response format.

The experiment thus followed a 2 (group: deductive vs. probabilistic reasoning) x 3 (content: low conditional probability vs. high conditional probability vs. abstract content) x 2 (validity: valid vs. invalid) mixed design, with "group" as a between-subjects variable and "content" and "validity" as within-subjects variables. As dependent variables we measured people's binary or graded responses and the neural activations during presentation of the second premise and the evaluation of the conclusion (cf. Rodrigo-Moreno & Hirsch, 2009; Luo et al., 2014; Prado et al., 2010).

5.3. *Procedure*

The experiment was conducted in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Before entering the scanner, participants were familiarized with the task on a desktop computer. They were told that they would see problems consisting of an if–then statement, a factual statement, and a conclusion that they would have to evaluate. Half of the participants were carefully instructed to evaluate the conclusions deductively, and the other half, probabilistically. This was done twice, at the beginning and at the end of the instruction part. In eight practice problems, participants learned to handle the flexible response format. They could only participate in the fMRI experiment when they were able to describe their instructions in their own words. Instructions and four practice problems were again presented at the beginning of the scanning session.

The 96 conditional inference problems were presented randomly in two separate runs of 48 problems each. The baseline problems were presented randomly in a third run, at the end of the experiment. A pause was possible between runs. Both premises and the conclusion were presented on separate screens for a fixed time interval (see Fig. 2). Conclusions were presented in red font together with the flexible response format. Between the problems, a jittered ITI of 1500–3750 msec was used.

5.4. *fMRI acquisition*

Magnetic resonance images were collected on a 3 T whole-body scanner (Magnetom Prisma, Siemens Medical Systems, Erlangen, Germany) with a 64-channel head coil. For functional imaging, we acquired gradient-echo echo-planar images (EPI; TR = 2200 ms, TE = 30 ms, 40 slices, voxel size = 3 × 3 × 3 mm, gap = 0.75 mm, FoV = 216 × 216 × 150

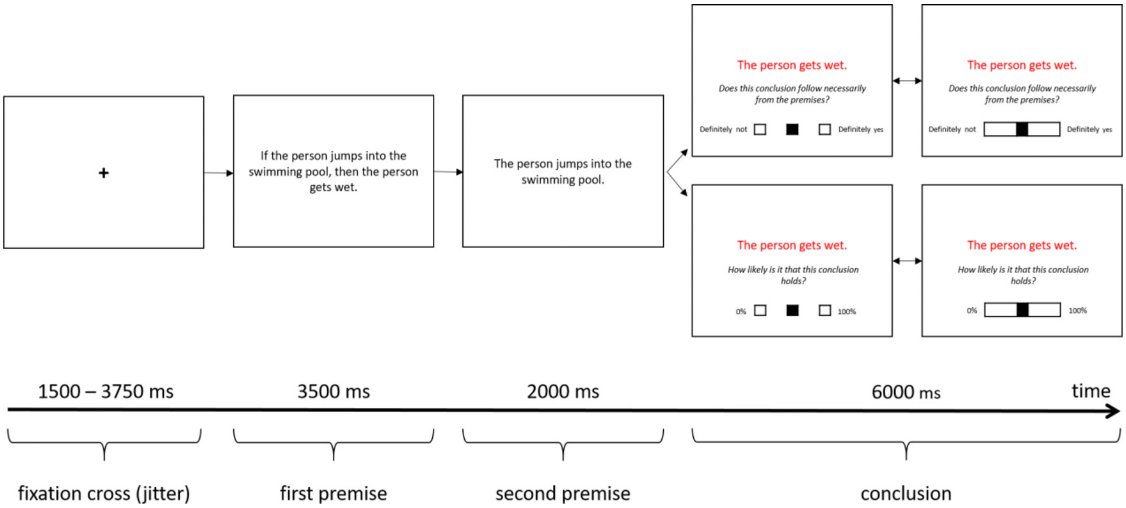

Fig. 2. Exemplar trial and stimuli presentation.

mm^3; flip angle = 75°, acceleration factor = 2) in a descending sequence with an orientation of -30° to the anterior commissure-posterior commissure line. Functional data were collected in three separate runs. During the first two runs, participants were confronted with the conditional inference problems. These two runs were of approx. 308 volumes. The third run contained the baseline problems and was of approx. 155 volumes. To correct for distortion of the functional images due to B_0 inhomogeneity, we also used a gradient-echo field map sequence (TE 1 = 10 ms, TE 2 = 12.46 ms, TR = 1000 ms, 40 slices, voxel size = 2 × 2 × 3 mm). Anatomical images were collected using a T1-weighted sequence (MPRAGE, TR = 1880 ms, TE = 3.53 ms, voxel size = 0.9 × 0.9 × 0.9 mm). Head movements were decreased by the use of foam pads and scanner noise was minimized with ear plugs.

5.5. fMRI analysis

Preprocessing and data analysis were performed using the software package Statistical Parametric Mapping implemented in Matlab (SPM12; Wellcome Trust Centre for Neuroimaging, London, UK). Raw functional data quality was checked with the ArtRepair toolbox (Mazaika, Hoeft, Glover, & Reiss, 2009) and each detected bad volume was later included into the statistical model as a separate regressor in the first-level analysis. The following preprocessing steps were applied: reorientation; realignment and unwarping, with realignment parameters examined on a volume-to-volume basis and cut-off values set at 1 mm for translation and 0.5° for rotation; slice timing correction; normalization to MNI space based on segmentation of the anatomical image and with resampling to 2 × 2 × 2 mm; and smoothing using an 8 mm Gaussian kernel. Data were high-pass filtered (cut-off: 128 s) and whitened using an AR(1) model.

5.6. fMRI models

For the statistical analysis, trial-related activity for each participant was modelled by a vector of trial onsets convolved with a canonical hemodynamic response function. For each reasoning group (deductive and probabilistic) we set up two models: one model to measure neural processes at the time of the presentation of the second premise (henceforth: *P2-model*), and a second model to measure neural processes at the time of the presentation of the conclusion (henceforth: *C-model*). For the first two runs containing the inference problems, we used the same regressors in both models: three boxcar regressors captured activity related to the presentation of the first premise (first premise with low conditional probability, high conditional probability, and abstract content); six boxcar regressors were specified to represent the presentation of the second premise (the second premise was subdivided into the same categories as the first, and additionally each regressor was split by MP and MT); and twelve boxcar regressors captured activity related to the presentation of the conclusion (which was also differentiated into low and high conditional probability and abstract content, and each of these partitions was split by MP, MT, invalid MP, and invalid MT). However, for the baseline (the third run), the two models differed in the specification of their regressors. For this run, different regressors were specified for each model in order to avoid multicollinearity. In both models, one boxcar regressor was specified to represent the presentation of the first premise baseline items. In the P2-model, three boxcar regressors represented the presentation of the second premise (low conditional probability, high conditional probability, abstract content) and two boxcar regressors characterized the presentation of the conclusion (valid/invalid). In the C-model it was exactly the opposite: three boxcar regressors captured activity related to the presentation of the conclusion (low conditional probability, high conditional probability, abstract content) and two boxcar regressors represented the presentation of the second premise (valid/invalid). All regressors were set to the respective presentation durations. Efficiency estimations in FEAT (FMRIB software library; Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012) prior to the analysis showed that this design was superior to alternative designs with varying jitters and varying trial combinations. A general linear model (GLM) was calculated for each participant to model the effects of interest and six regressors capturing residual motion-related artefacts.

5.7. fMRI contrasts

We set up two types of contrasts for the analysis of our data: (1) contrasts that allowed us to test the activation patterns for either deductive or probabilistic reasoning alone, and (2) contrasts for the direct assessment of differences between the two kinds of reasoning. For the analysis of the activation patterns of deductive and probabilistic reasoning alone, we set up for each model (the P2-model and the C-model) the following three t-contrasts for content (high conditional probability/low conditional probability/abstract) on first level, followed by a random-effects second-level analysis to assess group effects: overall (high conditional probability + low conditional probability + abstract vs. baseline), high (high conditional probability vs. baseline), low (low conditional probability vs. baseline), and abstract problems (abstract vs. baseline). For the analysis of the differences between deductive and probabilistic reasoning we set up, for each model (the P2-model and the C-model), t-contrasts for content with the same differentiation on first level: overall

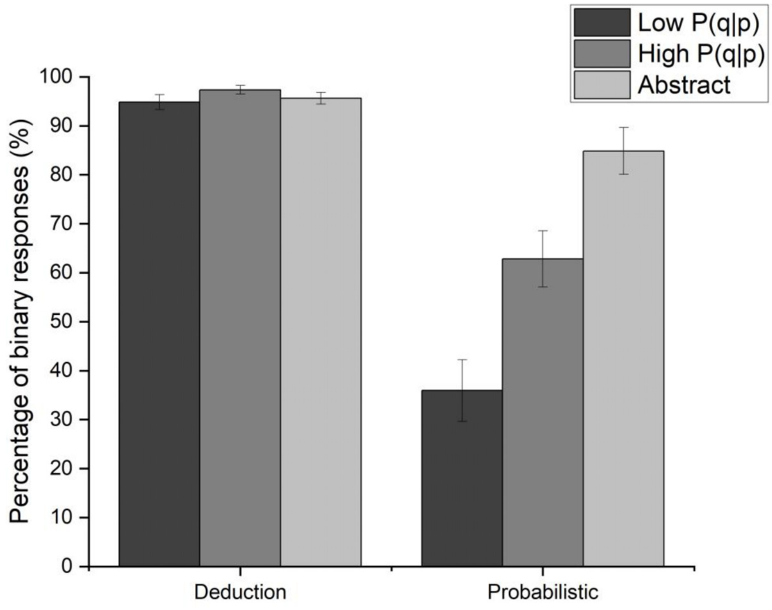

Fig. 3. Percentage of binary responses (%) for problems of low conditional probability, high conditional probability, and abstract content in the deductive and probabilistic reasoning groups. Error bars show standard errors.

(high conditional probability + low conditional probability + abstract vs. baseline), high (high conditional probability vs. baseline), low (low conditional probability vs. baseline), and abstract problems (abstract vs. baseline). Then, we employed t-tests for independent measures (two-sample t-tests) on second level.

Brain regions for analysis were determined based on their involvement in memory retrieval in probabilistic reasoning, and inhibition of prior knowledge, cognitive control, and conflict monitoring in deductive reasoning. Results were initially thresholded at $p < .0005$ (uncorrected) and corrected for multiple comparisons using small volume correction (SVC) in a priori regions of interest (ROIs; $p < .05$, family-wise error rate [FWE] corrected). Accordingly, for display purposes, all SVC significant activations are displayed at the initial threshold. SVC was carried out using anatomical masks of bilateral Hippocampus, LPFC, and ACC from the Harvard-Oxford probabilistic atlas (Harvard Center for Morphometric Analysis). Apart from our ROIs, data were analysed using a whole-brain FWE-corrected threshold ($p < .05$). Activations are displayed overlaid onto the averaged structural MRI scan of study participants. Stereotaxic coordinates are given in MNI space.

6. Results

6.1. Preferred response format

We first analyzed how often participants decided to give a binary response in the inference task.[2] Descriptive data can be found in Fig. 3.

A 2 (group: deductive vs. probabilistic reasoning) x 3 (content: low conditional probability vs. high conditional probability vs. abstract) mixed Analysis of Variance (ANOVA) on the percentage of binary responses revealed main effects for group, $F(1,\ 38) = 50.86$, $p < .001$, $\eta_p^2 = .572$, and content, $F(1.53,\ 58.05) = 42.85$, $p < .001$, $\eta_p^2 = .530$, but also an interaction between both factors, $F(1.53,\ 58.05) = 39.76$, $p < .001$, $\eta_p^2 = .511$. Therefore, we now concentrate only on the interaction. As expected, the content of inferences only affected participants in the probabilistic reasoning group, $F(1.48,\ 28.13) = 43.01$, $p < .001$, $\eta_p^2 = .694$. Participants gave more binary responses for inferences with abstract content ($M = 84.84\%$, $SD = 21.42$), followed by inferences with high conditional probability ($M = 62.81\%$, $SD = 25.71$; $t(19) = 3.92$, $p = .001$, $d = 0.925$[3]), followed by inferences with low conditional probability ($M = 35.94\%$, $SD = 28.22$; $t(19) = 7.73$, $p < .001$, $d = 0.983$) (Bonferroni-adjusted alpha: 0.025). In the deductive reasoning group, in contrast, the selection of binary responses was constantly high, not varying with content, $F(1.54,\ 29.27) = 2.65$, $p = .099$, $\eta_p^2 = .123$.

6.2. Acceptance ratings

We next analyzed participants' acceptance ratings for the conditional inference task. To make the participants responses comparable, binary and graded responses were mapped onto the same 7-point Likert scale by coding "definitely no" and "0%" responses as 1, and "definitely yes" and "100%" responses as 7. Values of 2–6 were used for answers in between. Descriptive statistics can be found in Fig. 4.

A 2 (group: deductive vs. probabilistic reasoning) x 3 (content: low conditional probability vs. high conditional probability vs. abstract) x 2 (validity: valid vs. invalid) ANOVA revealed a three-way interaction between instructions, validity, and content, $F(2,\ 76) = 19.07$, $p < .001$, $\eta_p^2 = .334$,[4] which allowed us to further analyze our data with two separate 2 (validity: valid vs. invalid) x 3 (content: low conditional probability vs. high conditional probability vs. abstract) ANOVAs, one for deductive and one for probabilistic instructions.

As expected, content only affected acceptance ratings in the probabilistic reasoning group, but not in the deductive reasoning group. For deductive reasoning, content had no effect on acceptance ratings; there was only a main effect of validity, $F(1,\ 19) = 3906.73$, $p < .001$, $\eta_p^2 = .995$, but no main effect of content or interaction with content (both F's < 1.39, p's > .262). For probabilistic reasoning, in contrast, content affected inferences. We found a main effect of validity, $F(1,\ 19) = 553.72$, $p < .001$, $\eta_p^2 = .967$, no main effect of content, $F(2,\ 38) = 1.34$, $p = .273$, $\eta_p^2 = .066$, but an interaction between validity and content, $F(2,\ 38) = 19.31$, $p < .001$, $\eta_p^2 = .504$. For valid inferences, acceptance ratings *increased* from problems with low conditional probability ($M = 5.87$, $SD = 0.58$) to problems with high conditional

---

[2] The preferred response for the baseline problems was, in both groups, a binary response (deduction group: 99.79%; probabilistic group: 93.75%).

[3] 3 Standardized mean differences (d) were computed as described by Borenstein (2009).

[4] In addition to the three-way interaction, the ANOVA also revealed a main effect of validity, $F(1,\ 38) = 2342.25$, $p < .001$, $\eta_p^2 = .984$, and interactions between validity and group, $F(1,\ 38) = 36.35$, $p < .001$, $\eta_p^2 = .489$, and between validity and content, $F(2,\ 76) = 12.71$, $p < .001$, $\eta_p^2 = .251$. All other effects were not significant, F's < 1.92, p's > .174.

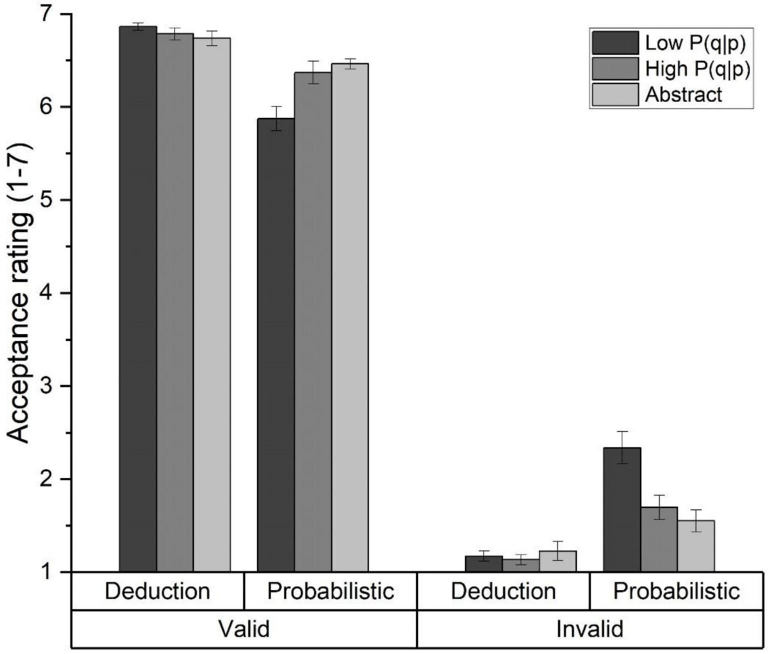

Fig. 4. Acceptance ratings (1–7) for valid and invalid problems of low conditional probability, high conditional probability, and abstract content in the deductive and probabilistic reasoning groups. Error bars show standard errors.

probability ($M = 6.37$, $SD = 0.55$; $t(19) = 4.48$, $p < .001$, $d = 0.877$), and – descriptively – to problems with abstract content ($M = 6.46$, $SD = 0.24$; $t(19) = 0.71$, $p = .486$, $d = 0.216$; Bonferroni-adjusted alpha: 0.025). This linear trend (low < high < abstract) was significant, $F(1, 19) = 19.32$, $p < .001$, $\eta_p^2 = .504$. For invalid inferences, in contrast, acceptance ratings *decreased* from problems with low conditional probability ($M = 2.34$, $SD = 0.78$) to problems with high conditional probability ($M = 1.70$, $SD = 0.57$; $t(19) = 6.00$, $p < .001$, $d = 0.859$), and – descriptively – to problems with abstract content ($M = 1.55$, $SD = 0.54$; $t(19) = 1.07$, $p = .296$, $d = 0.266$; Bonferroni-adjusted alpha: 0.025). This linear trend (low > high > abstract) was significant too, $F(1, 19) = 20.96$, $p < .001$, $\eta_p^2 = .524$.

6.3. fMRI Results

We now report the most relevant neural activations related to our hypotheses about the relation between deductive and probabilistic reasoning in the human brain. An overview of our complete results, including peak statistics and MNI coordinates, can be found in Tables 1, 2 and 3.[5] We start with the results for deductive reasoning, then give those for probabilistic reasoning, both compared to baseline, and then conclude with those for the two contrasts between the two reasoning conditions. As we are using three different ROIs (LPFC, ACC, and Hippocampus), we adjusted our alpha for the SVC results with Bonferroni Holmes. For LPFC our alpha was 0.017, for ACC it was 0.025 and for the Hippocampus it was 0.05.

6.4. Deductive reasoning

One of our main hypotheses was that deductive reasoning should be accompanied by cortical activity in regions known to be involved in mediating conflict and inhibition processes. Hence, we were particularly interested in patterns of activation in ACC and right LPFC, especially DLPFC and IFC (Goel et al., 2000; Prado & Noveck, 2007; Tsujii et al., 2010; Botvinick et al., 2004), which we explored on a whole-brain level as well as using two different ROIs (ACC and LPFC).

As expected, in the deductive reasoning group we found significant rACC activations during processing of the second premise for all problems compared to baseline (Fig. 5A; $p = .022$, $z = 4.05$, SVC). The same rACC activations were found for problems with low conditional probability compared to the baseline ($p = .046$, $z = 3.83$, SVC), and abstract problems compared to baseline ($p = .018$, $z = 4.11$, SVC), although the former did not reach the adjusted alpha level of 0.025. We also measured significant dACC activation during the processing of the second premise for all problems compared to baseline ($p = .001$, $z = 4.95$, SVC), for high-conditional-probability problems compared to baseline ($p = .020$, $z = 4.08$, SVC), low-conditional-probability problems compared to baseline ($p = .001$, $z = 4.89$, SVC), and abstract problems compared to baseline ($p < .001$, $z = 5.12$, SVC). Moreover, also in accordance with our predictions, in the deductive reasoning group, we found elevated activity in right IFC during the processing of the second premise for all problems compared to baseline (Fig. 5B; $p = .017$, $z = 4.68$, SVC) and for low-conditional-probability problems compared to baseline ($p = .041$, $z = 5.03$, whole-brain FWE). Importantly, all of these activations were only found in the deductive reasoning group, but not in the probabilistic reasoning group. These results support the conclusion that deductive reasoning is mediated by ACC and right IFC activity, suggesting the inhibition of prior knowledge in deduction.

We also predicted that deductive reasoning should lead to additional activity either in parietal areas, which would speak for the involvement of spatial representations and processes, or in temporal areas, which would speak for language-related processes in deductive reasoning. Our analysis only yielded additional parietal activations in the precuneus during the processing of the second premise for all problems compared to baseline (Fig. 5C, $p = .004$, $z = 5.47$, whole-brain FWE). This activation in precuneus was also found for low-conditional-probability problems compared to baseline ($p = .003$, $z = 5.51$, whole-brain FWE),

---

[5] Our computations include data from two participants for which we had reduced amounts of data points. For one participant, during the first run of the experiment, the scanner reported an error and stopped scanning until the end of this run. We lost approximately thirty seconds of data collection for this participant. Another participant stopped the experiment while performing the baseline task. We had to interrupt the experiment at that point and lost approximately one minute of data collection in that third run of the experiment. However, we did not exclude these participants from analysis because we still collected a sufficient amount of data from them.

Table 1
fMRI results for deductive reasoning.

Regions	Peak Statistics		
	p-value	z-score	MNI coordinates (x, y, z)
Whole-brain FWE corrected, p < .05			
Overall: High+Low+Abstract vs. Baseline			
Precuneus	.004	5.47	10, -66, 20
Occipital cortex	.001	5.70	-14, -72, 16
	.013	5.25	14, -72, 12
Somatosensory cortex	.001	5.68	56, -26, 22
	.012	5.33	-62, -22, 18
	.033	5.38	-58, -22, 38
Motor cortex	.005	5.44	-36, -16, 38
High vs. Baseline			
Precuneus	.044	5.02	10, -66, 18
Occipital cortex	.027	5.11	-14, -64, 14
Somatosensory cortex	<.001	5.90	56, -24, 24
	.020	5.11	-54, -28, 18
Low vs. Baseline			
Precuneus	.003	5.51	8, -68, 20
Occipital cortex	.003	5.51	-12, -70, 6
IFC	.041	5.03	54, 34, 2
mCC	.037	5.06	12, -18, 44
Somatosensory cortex	.002	5.62	-46, -24, 40
	.003	5.53	56, -28, 24
Abstract vs. Baseline			
Precuneus	.008	5.35	10, -66, 18
Occipital cortex	.002	5.59	-16, -72, 12
Somatosensory cortex	.002	5.61	-62, -20, 36
	.009	5.32	56, -16, 44
	.016	5.21	-64, -22, 16
	.015	5.23	-44, -30, 18
Motor cortex	.003	5.50	-36, -18, 40
*ROIs, SVC corrected, p < .05**			
Overall: High+Low+Abstract vs. Baseline			
IFC	.017	4.68	54, 34, 2
rACC	.022	4.05	6, 40, 2
dACC/ mCC	.001	4.95	-4, -10, 40
High vs. Baseline			
dACC/mCC	.020	4.08	-4, 2, 44
	.031	3.96	-4, -8, 38
Low vs. Baseline			
IFC	.003	5.03	54, 34, 2
rACC	.046	3.83	2, 40, -4
dACC/ mCC	.001	4.89	4, -4, 40
Abstract vs. Baseline			
rACC	.018	4.11	6, 40, 0
dACC/ mCC	.021	4.07	-4, 2, 44
dACC/ mCC	<.001	5.12	-4, -10, 40

All activations resulted from the P2-model. There were no significant results in the C-model. Peak statistics for all SVC-corrected activations initially thresholded at $p < .0005$ (uncorrected) and minimum cluster size threshold $k = 5$ voxels.

* p-values were tested against Bonferroni Holmes adjusted alphas (LPFC: alpha = 0.017; ACC: alpha = 0.025; Hippocampus: alpha = 0.05).

high-conditional-probability problems compared to baseline ($p = .044$, $z = 5.02$, whole-brain FWE), and abstract problems compared to baseline ($p = .008$, $z = 5.35$, whole-brain FWE). Parietal activations were found in the deductive reasoning group, but not in the probabilistic reasoning group. These results indicate that spatial representation and processing are particularly important for deductive reasoning.

6.5. *Probabilistic reasoning*

A further hypothesis was that probabilistic reasoning should lead to increased neural activation in the hippocampus because of its role in memory retrieval (e.g., Addis et al., 2007; Eldridge et al., 2000; Svoboda et al., 2006). In fact, we found increased hippocampal activity in the probabilistic reasoning group during the processing of the second premise for high-conditional-probability problems ($p = .030$, $z = 3.89$, SVC) and also during the processing of the conclusion for high-conditional-probability problems compared to baseline ($p = .032$, $z = 3.84$, SVC) and low-conditional-probability problems compared to baseline (Fig. 6A; $p = .046$, $z = 3.74$, SVC). This activity was also obtained for the processing of the conclusion for abstract problems compared to baseline ($p = .035$, $z = 3.79$, SVC). No hippocampal activations were found in the deductive reasoning group. These results indicate that memory retrieval plays an important role in probabilistic reasoning.

6.6. *Deductive versus probabilistic reasoning*

A further comparison is the direct contrast between deductive versus probabilistic reasoning. This analysis suggests higher activity in the IFC in the deductive reasoning group for abstract problems compared to baseline during the processing of the second premise (Fig. 5C; $p = .019$, $z = 4.59$, SVC)(adjusted alpha = 0.017). The IFC seems thus to be important in deductive but not probabilistic reasoning.

Table 2
fMRI results for probabilistic reasoning.

Regions	Peak Statistics		
	p-value	z-score	MNI coordinates (x, y, z)
Whole-brain FWE corrected, $p < .05$			
P2-Model			
Overall: High+Low+Abstract vs. Baseline			
Motor Cortex	.026	5.11	52, -2, 12
High vs. Baseline			
VMPFC	.015	5.22	-12, 44, -2
Abstract vs. Baseline			
Motor Cortex	.020	5.16	52, -2, 12
C-Modell			
Insula	.021	5.15	-38, -8, -8
*ROIs, SVC corrected, $p < .05$**			
P2-Model			
High vs. Baseline			
Hippocampus	.030	3.89	30, -12, -20
C-Model			
High vs. Baseline			
Hippocampus	.032	3.84	-24, -30, -12
Low vs. Baseline			
Hippocampus	.046	3.74	-26, -32, -12
Abstract vs. Baseline			
Hippocampus	.035	3.79	-22, -14, -16

Peak statistics for all SVC-corrected activations initially thresholded at $p < .0005$ (uncorrected) and minimum cluster size threshold $k = 5$ voxels.

* p-values were tested against Bonferroni Holmes adjusted alphas (LPFC: alpha = 0.017; ACC: alpha = 0.025; Hippocampus: alpha = 0.05).

Table 3
fMRI results for deductive vs. probabilistic reasoning.

Regions	Peak Statistics		
	p-value	z-score	MNI coordinates (x, y, z)
Deductive vs. probabilistic reasoning			
*ROIs, SVC corrected, $p < .05$**			
P2-Model			
Abstract vs. Baseline			
IFC	.019	4.59	54, 32, 2
Probabilistic vs. deductive reasoning			
*ROIs, SVC corrected, $p < .05$**			
C-Model			
Overall: High+Low+Abstract vs. Baseline			
Hippocampus	.041	3.69	-34, -18, -18
Abstract vs. Baseline			
Hippocampus	.011	4.07	-34, -18, -18

Peak statistics for all SVC-corrected activations $p < .05$) initially thresholded at $p < .0005$ (uncorrected) and minimum cluster size threshold $k = 5$ voxels.

* p-values were tested against Bonferroni Holmes adjusted alphas (LPFC: alpha = 0.017; ACC: alpha = 0.025; Hippocampus: alpha = 0.05).

6.7. *Probabilistic versus deductive reasoning*

Importantly, the reversed contrast – the direct comparison between probabilistic versus deductive reasoning – revealed significantly elevated activity in the hippocampus in the probabilistic reasoning group during conclusion processing. This higher hippocampus activation was obtained for all problems compared to baseline as well as for abstract problems compared to the baseline (Fig 6B; $p = .041$, $z = 3.69$; $p = .011$, $z = 4.07$, SVC, respectively). The hippocampus is thus important in probabilistic but not in deductive reasoning.

7. Discussion

When people reason in daily life they often consider their prior knowledge about the content of inferences. This has motivated many researchers to argue that everyday reasoning is uncertain and thus often based on probabilistic inferences (e.g., Evans & Over, 2004; Evans, 2012; Oaksford, 2015). Some researchers, however, go even further and argue that *all* human reasoning is probabilistic. But is probabilistic reasoning sufficient for human rationality? Isn't the ability to reason deductively, and even in new domains about which we do not have prior knowledge, a hallmark of human intelligence and rationality? We think so, and therefore we argued in this paper that human reasoning is not only probabilistic. Instead, we proposed that people can also reason deductively and that deductive and probabilistic reasoning are distinct processes with their own specific merits. Our assumptions are supported by our behavioral and neuroimaging findings.

On the behavioral level, we were able to show that content only affected inferences in the probabilistic but not in the deductive reasoning group. This indicates that – when reasoning with conditionals – people can indeed inhibit their prior knowledge and reason in a knowledge-free fashion. This inhibition of prior knowledge may certainly depend

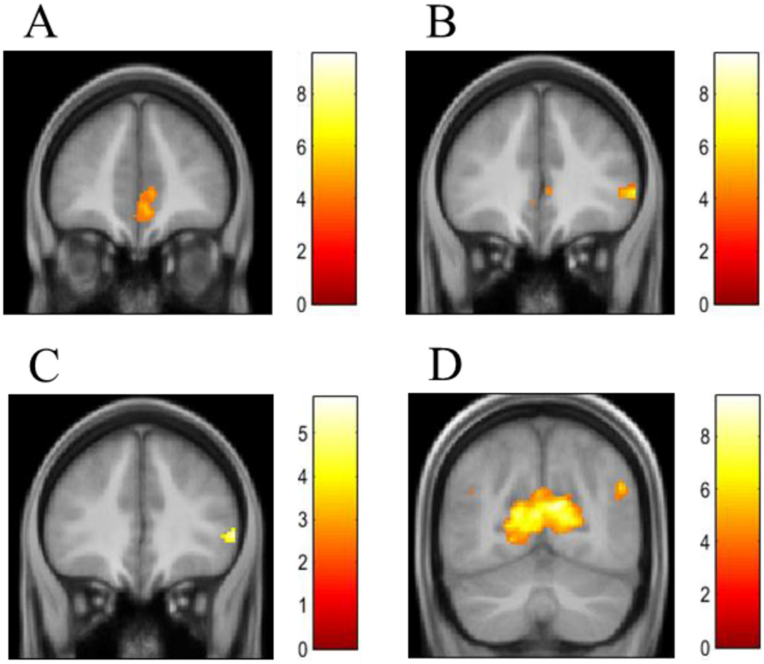

Fig. 5. ACC, IFC, and parietal activity mediate deductive reasoning (A–D). Significant activations for the deductive reasoning group during the processing of the second premise (P2-model) (SVC corrected, p < .05) in the (A) rACC (ACC ROI; contrast high + low + abstract vs. baseline (presented): 6, 40, 2; contrast abstract vs. baseline: 6, 40, 0), (B) IFC (LPFC ROI; contrast high + low + abstract vs. baseline (presented); contrast low vs. baseline: both 54, 34, 2), (C) IFC in the direct comparison between deductive vs. probabilistic reasoning (marginally significant)(contrast abstract vs. baseline (presented): 54, 32, 2), and (D) in the precuneus (whole-brain FWE corrected, p < .05; contrast high + low + abstract vs. baseline (presented): 10, -66, 20; contrast low vs. baseline: 8, -68, 20; contrast high vs. baseline; contrast abstract vs. baseline: both 10, -66, 18). Peak MNI coordinates are displayed. Clusters are shown in neurological orientation at a display threshold of p < .0005 (uncorrected), k = 5 voxels. Colour bars indicate t values. For clarity, the information whether the activations stem from SVC or FEW analyses are underlined.

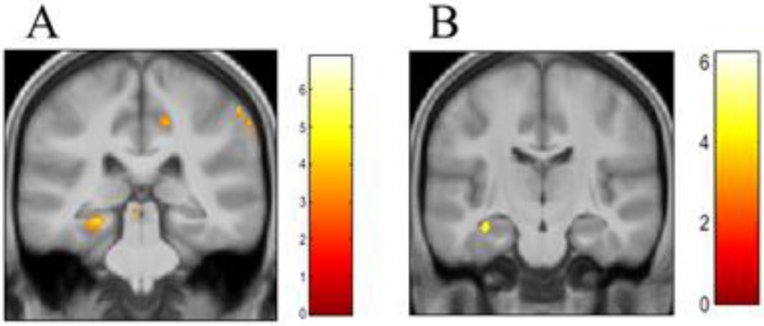

Fig. 6. Hippocampal activity underlies probabilistic reasoning (A–B). (A) Significant activations for the probabilistic reasoning group during the processing of the conclusion (C-model) in the hippocampus (hippocampus ROI; SVC corrected, p < .05)(contrast high vs. baseline: -24, -30, -12; contrast low vs. baseline (presented): -26, -32, -12; contrast abstract vs. baseline: -22, -14, -16), and (B) in the direct comparison between probabilistic vs. deductive reasoning (hippocampus ROI; SVC corrected, p < .05)(contrast high + low + abstract vs. baseline; contrast abstract vs. baseline (presented): -34, -18, -18). Peak MNI coordinates are displayed. Clusters are shown in neurological orientation at a display threshold of p < .0005 (uncorrected), k = 5 voxels. Colour bars indicate t values. For clarity, the information whether the activations stem from SVC or FEW analyses are underlined.

on how participants are instructed (Vadeboncoeur and Markovits, 1999) and on which kind of inference problems they have to solve (cf. Goel & Dolan, 2003). Yet, our findings nonetheless show that there are situations where people can reason deductively. Importantly, and somewhat unexpectedly, even in the probabilistic group there was still a considerable number of participants that preferred the dichotomous response format and gave binary responses. This was the case for abstract conditionals (85%), conditionals with high conditional probability (63%), and even for conditionals with low conditional probability (36%). As a result, valid inferences were accepted more strongly for abstract conditionals than for conditionals with low conditional probabilities, and invalid inferences more strongly for low-conditional-probability problems than for abstract conditionals.

It is interesting that so many participants in the probabilistic group decided to give binary responses. One of the main arguments of probabilistic theories is that people want to use their prior knowledge during reasoning and that this is what leads to probabilistic and uncertain reasoning. This was actually the reason why our flexible response format always started with the dichotomous format that had to be extended into the scale. We used this order because a switch to the scale would indeed show that people feel the urge to give a graded response, as suggested by probabilistic theory. Yet, even in the probabilistic reasoning group only a few participants expanded the dichotomous response format into a scale. The easiest explanation, of course, is that this is just an artifact of the order of response formats. However, we do not think that this explanation is sufficient. As just explained, probabilistic theorists argue that people always want to give graded responses. They argue that people's beliefs are represented in "many shades of grey" and not in "black" or "white" as in binary logic (Evans, 2012, p. 9; cf. Oaksford & Chater, 2007; Over, 2009). Our finding that so many participants did not switch to the scale challenges this assumption. If people really had such a strong impulse to give a graded response, that should not be prevented by the very small extra effort needed to switch to the scale format. This is an important empirical finding on the behavioural level: it shows that participants do not always feel the desire to consider degrees of belief and to reason probabilistically. In fact, they also often want to give a clear "yes" or "no" answer.

This is also supported by the neuroimaging findings, which also speak for a coexistence of deductive and probabilistic reasoning in the human brain. In the probabilistic reasoning group, but not in the deductive group, we found – as expected – activation in the hippocampus related to memory retrieval (Eldridge et al., 2000; Greicius et al., 2003). In contrast, in the deductive group, but not in the probabilistic group, we found activations in the ACC and the IFC – both regions that are related to conflict monitoring and inhibition (e.g., Botvinick et al., 2001, 2004; Tsujii et al., 2010; Aron et al., 2003). In addition, in this group we also found activations in parietal regions that have been related to spatial representation and deductive reasoning (e.g., Goel & Dolan, 2001; Knauff et al., 2002, 2003; Fangmeier et al., 2006; Mackey et al., 2013;

Knauff, 2013). Again, we did not find these activations during probabilistic reasoning.

The reported activation pattern suggests that prior knowledge is activated during both deductive and probabilistic reasoning, but that the way this prior knowledge is handled differs between the two kinds of reasoning. In probabilistic reasoning, prior knowledge is essential for the evaluation of conclusions. Reasoners have to retrieve their knowledge about the content of an inference and decide on the basis of this knowledge the probability of the conclusion. This explains the hippocampal activation. In deductive reasoning, prior knowledge is also retrieved, but subsequently inhibited in order to be able to reason deductively. Therefore, no hippocampal activations are observable, but only the activations related to the inhibition of prior knowledge. However, once prior knowledge is inhibited, reasoning proceeds in a knowledge-free way that is qualitatively different from probabilistic reasoning and seems to rely on spatial representations and processes – as indicated by the parietal activation unique for deductive reasoning. This finding supports the mental models theory of deductive reasoning (e.g., Knauff et al., 2002, 2003; Fangmeier et al., 2006).

We admit that there might be alternative interpretations for the IFC activation. For instance, one reviewer suggested that if the right IFC is more active during deductive vs. probabilistic reasoning in the abstract problems, then the inhibition interpretation would be questionable. The right IFC could in fact be active regardless of response inhibition (Hampshire et al., 2010) and rule representations might be decoded from its activity patterns (Reverberi et al., 2012). The activation we found in the right IFC could thus result from higher attentional control instead of response inhibition per se (cf. Hampshire et al., 2010). We cannot completely rule out this interpretation, but consider it unlikely. First, because although our abstract problems were abstract with relation to the conditional relation between p and q (e.g., “If the box is empty, then the box has stars on it”), they still contained familiar terms about which the participants had prior knowledge (e.g., “box”, “stars” etc.; see Gazzo Castañeda & Knauff, 2021). That is, also for abstract problems participants had to inhibit prior knowledge to reason deductively – albeit not for the evaluation of the conclusion, but for decoding the logical form from the premises. The IFC activation is thus not surprising. And second, because we found a distinct activation pattern in the comparisons against baseline. Here, the ACC was active for all kinds of problems (abstract, high and low conditional probability), while the IFC was only activated during problems of low conditional probability, i.e., when participants are processing problems with many disabler and thus need most inhibition of prior knowledge. This is difficult to reconcile with the alternative explanation and supports our initial assumption that the IFC is important for inhibition of prior knowledge.

A final observation that supports our approach of distinguishing deductive and probabilistic reasoning processes is that we found all increased activations for deductive reasoning during the processing of the second premise, whereas most relevant activations for probabilistic reasoning appeared during conclusion processing. The higher activations during the processing of the second premise in deductive reasoning agrees with many previous findings (e.g., Rodriguez-Moreno & Hirsch, 2009; Luo et al., 2014; Prado et al., 2010). When reasoning deductively, most of the mental work can already be done when the second premise is presented. People then already can integrate the information from the two premises, which is the core of the deductive reasoning process. Therefore, when confronted with the to-be-evaluated conclusion, they can compare it with the conclusion they drew beforehand. In Fangmeier et al. (2006) we presented an fMRI study that carefully distinguished between the different phases of a reasoning process, and also identified this phase as essential for deductive inferences. However, the fact that for probabilistic reasoning we found significant activity mostly during conclusion processing shows that, here, people are reasoning differently. When people reasoned probabilistically, they waited for the conclusion and only then considered their prior knowledge to give a response. This is also reflected in our observation that most of the participants in the probabilistic reasoning group moved the cursor back and forth several times along the scale until they selected a response, whereas participants in the deductive group right away selected their “yes” or “no” response.

Our results draw a clear picture about human reasoning: people can reason both deductively and probabilistically, and the two are distinct cognitive processes. This reminds us of dual-process theories that postulate two modes of reasoning (Evans & Stanovich, 2013; Markovits, Brunet, Thompson, & Brisson, 2013). One mode is usually described as contextualized, automatic, and effortless, the other as abstract, controlled, and dependent on working memory (e.g., Evans & Stanovich, 2013; Evans, 2008). In particular the distinction between abstract and contextualized reasoning can be related to the distinction between deductive and probabilistic reasoning. Verschueren et al. (2005), for instance, showed in a behavioral study that people can solve inference problems in two different ways. They can do so probabilistically, by considering likelihood information, and more deductively, based on mental models. In this latter form of reasoning, conclusions do not depend on probability information, but on reasoners’ knowledge about counterexamples to the conclusion, i.e., situations in which the premises are true but the conclusion is not. This is also supported by the work of Markovits and colleagues, who could show that people can indeed reason either probabilistically or based on mental models (Markovits, Brisson, & de Chantal, 2017; Markovits, Brunet, Thompson, & Brisson, 2013; Markovits, de Chantal, Brisson, & Gagnon-St-Pierre, 2019; see also Gazzo Castañeda & Knauff, 2021). Similarly, Rips (2001) and Rotello and Heit (2009) have found evidence in favor of two distinct ways of thinking. Although they contrasted deductive and inductive reasoning rather than deductive and probabilistic reasoning, their experiments do show that people reason differently when they have to make evaluations on the necessity or the plausibility of arguments. Reasoning about the necessity of arguments depended mostly on analytic processing, while reasoning about their plausibility depended mostly on heuristic processing (Heit & Rotello, 2010). The present brain-imaging findings are novel, as they demonstrate for the first time that deductive and probabilistic reasoning processes also coexist on the brain level.

Proponents of probabilistic theories might nonetheless argue that the different activation patterns are just due to our baseline condition. The purpose of baselines in reasoning experiments is to control for reasoning-irrelevant activations, such as reading and text comprehension. Therefore, such baselines often have the same content as the actual inference problems, but without asking participants to reason. The activations found during the baseline problems are then subtracted from the activations found during the reasoning task, so that the “core” neural correlates for reasoning can be found. Oaksford (2015), however, argues that in this way the effect of prior knowledge disappears, because the baseline, which has the same content, is subtracted from the reasoning condition. As a result, he argues, many imaging studies fail to find prior-knowledge-related activations that were actually there. This criticism may also apply to our study, in which the baseline problems also had the same content as the inferences. So, maybe our baseline eliminated some retrieval-specific activations during deduction. The crucial point, however, is that we nevertheless found hippocampal activation in the probabilistic condition and conflict-monitoring and inhibition-related activity during deductive reasoning. Prior knowledge thus seems only to be necessary for probabilistic reasoning, where we were still able to find hippocampal activation (although we used a content-rich baseline), but not for deductive reasoning.

Advocates of probabilistic theories might also criticize that we did not distinguish between the response options chosen by participants. We found that deductive reasoning is accompanied by parietal, ACC, and IFC activations, and that probabilistic reasoning is related to activations in the hippocampus. A difficulty, though, is that we do not know whether these activations differed when participants either chose the binary or the scaled response option. To test this, we conducted a re-analysis with only those problems in which participants in the deduction group gave a binary response and those problems in which participants in the probability group gave a scaled response. For the deduction group, such an analysis was possible and we could replicate the activity in parietal, ACC, and IFC areas. For the probability group, however, we had too few observations to conduct this analysis, since participants did not give enough graded responses. We will explore this in future experiments. Yet, the problem of too few observations may remain, even with larger samples, because even under probabilistic instructions participants often give binary responses, as we have shown.

Finally, another criticism may be that our interpretations rely on reverse inferences, i.e., interpreting brain activities as indicators for certain kinds of cognitive processes. However, since we were aware of this problem from the start, we strictly avoided post-hoc cognitive interpretations of neural activations. We had clear cognitive hypotheses which we translated into clear hypotheses about neural activities. Only these hypotheses were tested and our study led to novel findings that are important for the psychology of human reasoning. Moreover, proponents of probabilistic theories might also refer to our relatively small sample size to question our results. But that is no convincing argument, as behavioral studies on conditional reasoning have shown large effects of instructions ($d = 0.96$; Singmann and Klauer, 2011, Exp. 1, MP-Inferences) and content in everyday reasoning (d = 0.82 [d_z=0.70]; Gazzo Castañeda & Knauff, 2021, Exp. 2, high vs. low conditional probability). Power analyses based on these effect sizes (using G*Power, Faul et al., 2009) show that our sample size is strong enough to detect such effects of instructions and content (80% power, alpha = .05, two tailed, with independent samples t-test and depended samples t-test, respectively). In fact, other fMRI studies on reasoning have used similar sample sizes (e.g., 20 participants in Coetzee & Monti, 2018; 16 in Luo et al., 2014; and 17 participants in Prado et al., 2010).

Some cognitive reasoning researchers have argued that most of the information we get in our daily life is only true up to a certain degree. They also emphasize that people consider their prior knowledge to evaluate this uncertain information. We do not deny either. However, they throw the baby out with the bathwater when they say that *all* reasoning is probabilistic. The ability to reason deductively is a hallmark of human intelligence and rationality, even though, of course, we also sometimes use our prior knowledge to solve inference problems. The present research provides new evidence for the coexistence of deductive and probabilistic processes in human reasoning. This has important implications for the psychology of reasoning. In recent years, the deductive and the probabilistic paradigm have become increasingly separated. The probabilistic paradigm claims to be “the new psychology of reasoning” and considers deductive reasoning to be the “old paradigm” (e.g., Evans, 2012; Elqayam & Over, 2013). This alleged opposition between “old” and “new”, however, hinders progress in the field (Knauff & Gazzo Castañeda, 2021). Our study shows that *both* paradigms are crucial for our empirical and theoretical understanding of human reasoning. Therefore, instead of promoting rivalry, cognitive psychologists should work harder on finding further connections between the two paradigms.

Credit author statement

Lupita Estefania Gazzo Castañeda: Conceptualization, Methodology, Formal analysis (behavioral), Investigation, Writing - original draft, Writing - Review & Editing, Visualization (behavioral) **Benjamin Sklarek:** Conceptualization, Methodology, Investigation, Writing - original draft **Dennis E. Dal Mas:** Methodology, Software, Formal analysis (fMRI), Investigation, Data Curation, Writing - original draft, Visualization (fMRI) **Markus Knauff:** Conceptualization, Methodology, Writing - original draft, Writing - Review & Editing, Recourses, Supervision, Project administration, Funding acquisition

Data and code availability statement

Due to privacy issues, the data for the experiment is available from the corresponding author on request and requires a DUA and an ethics approval from the other party’s institution. Our behavioral analyses were conducted with the codes implemented in IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA). The code used for fMRI analyses was retrieved from the statistical parametric mapping (SPM) platform.

Funding

This research was supported by DFG grant KN 465/9-2 and KN 465/10-2 within the Priority Program “New Frameworks of Rationality” (SPP 1516) to Markus Knauff. The funding institution was not involved in study design; collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

Acknowledgments

We thank Vinod Goel and Manuela Sellitto for their comments on our fMRI results.

References

Addis, D.R., Wong, A.T., Schacter, D.L., 2007. Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. Neuropsycholgia 45, 1363–1377. doi:10.1016/j.neuropsychologia.2006.10.016.

Aron, A., Fletcher, P.C., Bullmore, E.T., Sahakian, B.J., Robbins, T.W., 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat. Neurosci. 6, 115–116. doi:10.1038/nn1003.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London,* 53, 370–418. (Reprinted in E. S. Pearson & M. G. Kendall (Eds.), *Studies in the History of Statistics and Probability*. London, England: Griffin, 1970)

Boole, G., 1854. An Investigation of The Laws of Thought on Which are Founded the Mathematical Theories of Logic And Probabilities. Macmillan, London, England.

Borenstein, M., 2009. Effect sizes for continuous data. In: Cooper, H., Hedges, L.V., Valentine, Jeffrey C. (Eds.), The Handbook of Research Synthesis and Meta-Analysis. Russell Sage Foundation, New York, NY, pp. 221–235.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol. Rev. 108, 624–652. doi:10.1037/0033-295X.108.3.624.

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8, 539–546. doi:10.1016/j.tics.2004.10.003.

Braine, M.D.S., 1978. On the relation between the natural logic of reasoning and standard logic. Psychol. Rev. 85, 1–21. doi:10.1037/0033-295X.85.1.1.

Braine, M.D.S., O’Brien, D.P, 1991. A theory of if: A lexical entry, reasoning program, and pragmatic principles. Psychol. Rev. 98, 182–203. doi:10.1037/0033-295X.98.2.182.

Braine, M.D.S., O’Brien, D.P, 1998. Mental Logic. Erlbaum, Mahwah, NJ.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. H. Bower (Ed.), *The psychology of Learning and Motivation* (Vol. 18, pp. 313–371). New York, NY: Academic Press.

Byrne, R.M.J., 1989. Suppressing valid inferences with conditionals. Cognition 31, 61–83. doi:10.1016/0010-0277(89)90018-8.

Coetzee, J.P., Monti, M.M., 2018. At the core of reasoning: Dissociating deductive and non-deductive load. Hum. Brain Mapp. 39, 1850–1861. doi:10.1002/hbm.23979.

Cummins, D.D., 1995. Naive theories and causal deduction. Mem. Cognit. 23, 646–658. doi:10.3758/BF03197265.

Cummins, D.D., Lubart, T., Alksnis, O., Rist, R., 1991. Conditional reasoning and causation. Mem. Cognit. 19, 274–282. doi:10.3758/BF03211151.

De Neys, W., Schaeken, W., d’Ydewalle, G., 2002. Causal conditional reasoning and semantic memory retrieval: a test of the semantic memory framework. Mem. Cognit. 30, 908–920. doi:10.3758/BF03195776.

De Neys, W., Schaeken, W., d’Ydewalle, G., 2003. Inference suppression and semantic memory retrieval: Every counterexample counts. Mem. Cognit. 31, 581–595. doi:10.3758/BF03196099.

De Neys, W., Vartanian, O., Goel, V., 2008. Smarter than we think: when our brains detect that we are biased. Psychol. Sci. 19, 483–489. doi:10.1111/j.1467-9280.2008.02113.x.

Eldridge, L.L., Knowlton, B.J., Furmanski, C.S., Bookheimer, S.Y., Engel, S.A., 2000. Remembering episodes: a selective role of the hippocampus during retrieval. Nat. Neurosci. 3, 1149–1152. doi:10.1038/80671.

Elqayam, S., Bonnefon, J.-F., Over, D.E., 2016. New Paradigm Psychology of Reasoning. Routledge, Abingdon, England (Originally published as a special issue of *Thinking & Reasoning*, 2013).

Elqayam, S., Over, D.E., 2013. New paradigm psychology of reasoning: an introduction to the special issue edited by Elqayam, Bonnefon, and Over. Think. Reason. 19, 249–265. doi:10.1080/13546783.2013.841591.

Evans, J.St.B.T, 2008. Dual-processing accounts of reasoning, judgment, and social cognition. Annu. Rev. Psychol. 59, 255–278.

Evans, J.St.B.T, 2012. Questions and challenges for the new psychology of reasoning. Think. Reason. 18, 5–31. doi:10.1080/13546783.2011.637674.

Evans, J.St.B.T., Handley, S.J., Over, D.E, 2003. Conditionals and conditional probability. J. Exp. Psychol.: Learn., Mem. Cognit. 29, 321–355. doi:10.1037/0278-7393.29.2.321.

Evans, J.St.B.T., Newstead, S.E., Byrne, R.M.J, 1993. Human Reasoning: the Psychology of Deduction. Erlbaum, Hove, England.

Evans, J.St.B.T., Over, D.E, 2004. If. Oxford University Press, Oxford, England.

Evans, J.St.B.T., Stanovich, K.E, 2013. Dual-process theories of higher cognition: advancing the debate. Perspect. Psychol. Sci. 8, 223–241.

Fangmeier, T., Knauff, M., Ruff, C.C., Sloutsky, V., 2006. fMRI evidence for a three-stage model of deductive reasoning. J. Cogn. Neurosci. 18, 320–334. doi:10.1162/jocn.2006.18.3.320.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav. Res. Methods 41, 1149–1160.

Frege, G., 1879. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Verlag von Louis Nebert, Halle, Germany.

Gazzo Castañeda, L.E., Knauff, M., 2013. Individual differences, imagery and the visual impedance effect. In: Knauff, M., Pauen, M., Sebanz, N., Wachsmuth, I. (Eds.), Proceedings of the 35th Annual Conference of the Cognitive Science Society. Cognitive Science Society, Austin, TX, pp. 2374–2379.

Gazzo Castañeda, L.E., Knauff, M., 2018. Quantifying disablers in reasoning with universal and existential rules. Think. Reason. 24, 344–365. doi:10.1080/13546783.2017.1401000.

Gazzo Castañeda, L.E., Knauff, M., 2021. Everyday reasoning with unfamiliar conditionals. Think. Reason. 27, 389–416. doi:10.1080/13546783.2020.1823478.

Goel, V., Buchel, C., Frith, C., Dolan, R.J., 2000. Dissociation of mechanisms underlying syllogistic reasoning. Neuroimage 12, 504–514. doi:10.1006/nimg.2000.0636.

Goel, V., Dolan, R.J., 2001. Functional neuroanatomy of three-term relational reasoning. Neuropsychologia 39, 901–909. doi:10.1016/S0028-3932(01)00024-0.

Goel, V., Dolan, R.J., 2003. Explaining modulation of reasoning by belief. Cognition 87, B11–B22. doi:10.1016/S0010-0277(02)00185-3.

Goel, V., Dolan, R.J., 2004. Differential involvement of left prefrontal cortex in inductive and deductive reasoning. Cognition 93, B109–B121. doi:10.1016/j.cognition.2004.03.001.

Goel, V., Gold, B., Kapur, S., Houle, S., 1997. The seats of reason? An imaging study of deductive and inductive reasoning. Neuroreport 8, 1305–1319. doi:10.1097/00001756-199703240-00049.

Greicius, M.D., Krasnow, B., Boyett-Anderson, J.M., Eliez, S., Schatzberg, A.F., Reiss, A.L., Menon, V., 2003. Regional analysis of hippocampal activation during memory encoding and retrieval: fMRI study. Hippocampus 13, 164–174. doi:10.1002/hipo.10064.

Hampshire, A., Chamberlain, S.R., Monti, M.M., Duncan, J., Owen, A.M., 2010. The role of the right inferior frontal gyrus: Inhibition and attentional control. Neuroimage 50, 1313–1319.

Heit, E., Rotello, C.M., 2010. Relations between inductive reasoning and deductive reasoning. J. Exp. Psychol.: Learn., Mem. Cognit. 36, 805–812. doi:10.1037/a0018784.

Houdé, O., Borst, G., 2015. Evidence for an inhibitory-control theory of the reasoning brain. Front. Neurosci. 9, 148. doi:10.3389/fnhum.2015.00148.

Houdé, O., Zago, L., Crivello, F., Moutier, S., Pineau, A., Mazoyer, B., et al., 2001. Access to deductive logic depends on a right ventromedial prefrontal area devoted to emotion and feeling. Neuroimage 14, 1486–1492.

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., et al., 2000. Shifting from the perceptual brain to the logical brain. J. Cognit. Neurosc. 12, 721–728.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. doi:10.1016/j.neuroimage.2011.09.015.

Johnson-Laird, P.N., 2001. Mental models and deduction. Trends Cogn. Sci. 5, 434–442. doi:10.1016/S1364-6613(00)01751-4.

Johnson-Laird, P.N., Byrne, R.M.J., 1991. Deduction. Erlbaum, Hove, England.

Johnson-Laird, P.N., Byrne, R.M.J., 2002. Conditionals: a theory of meaning, pragmatics, and inference. Psychol. Rev. 109, 636–678. doi:10.1037/0033-295X.109.4.646.

Johnson-Laird, P.N., Khemlani, S.S., Goodwin, G.P., 2015. Logic, probability, and human reasoning. Trends Cognit. Sci. 19, 201–214. doi:10.1016/j.tics.2015.02.006.

Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S., Caverni, J.-P., 1999. Naive probability: a mental model theory of extensional reasoning. Psychol. Rev. 106, 62–88. doi:10.1037/0033-295X.106.1.62.

Knauff, M., 2006. Deduktion und logisches Denken. In: Funke, J. (Ed.), Enzyklopädie der Psychologie, Themenbereich C, Bd. 8, Denken und Problemlösen. Hogrefe, Göttingen, Germany, pp. 167–264.

Knauff, M., 2009. A neuro-cognitive theory of deductive relational reasoning with mental models and visual images. Spatial Cognit. Comput. 9, 109–137. doi:10.1080/13875860902887605.

Knauff, M., 2013. Space to Reason: A Spatial Theory of Human Thought. MIT Press, Cambridge, MA.

Knauff, M., Gazzo Castañeda, L.E., 2021. When nomenclature matters: is the “new paradigm” really a new paradigm for the psychology of reasoning? Think. Reason. doi:10.1080/13546783.2021.1990126.

Knauff, M., Fangmeier, T., Ruff, C.C., Johnson-Laird, P.N., 2003. Reasoning, models, and images: behavioral measures and cortical activity. J. Cogn. Neurosci. 15, 559–573. doi:10.1162/089892903321662949.

Knauff, M., Johnson-Laird, P.N., 2002. Visual imagery can impede reasoning. Mem. Cognit. 30, 363–371. doi:10.3758/BF03194937.

Knauff, M., May, E., 2006. Mental imagery, reasoning, and blindness. Q. J. Exp. Psychol. 59, 161–177. doi:10.1080/17470210500149992.

Knauff, M., Mulack, T., Kassubek, J., Salih, H.R., Greenlee, M., 2002. Spatial imagery in deductive reasoning: a functional MRI study. Cognit. Brain Res. 13, 203–212. doi:10.1016/S0926-6410(01)00116-1.

Knauff, M., Spohn, W. (Eds.), 2021. *The Handbook of Rationality*. MIT Press, Cambridge, MA.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. Science 314, 829–832. doi:10.1126/science.1129156.

Kochari, A., van Rooij, R., Schulz, K., et al., 2020. Experimenting with subject alternatives for analysing generic sentences. In: Franke, M., et al. (Eds.), Proceedings of *Sinn und Bedeutung 24*, 1. Osnabrück University, pp. 411–428.

Lepage, M., Habib, R., Tulving, E., 1998. Hippocampal PET activations of memory encoding and retrieval: The HIPER model. Hippocampus 8, 313–322. doi:10.1002/(SICI)1098-1063(1998)8:4<313::AID-HIPO1>3.0.CO;2-I.

Luo, J., Tang, X., Zhang, E., Stupple, E.J.N., 2014. The neural correlates of belief-bias inhibition: The impact of logic training. Biol. Psychol. 103, 276–282. doi:10.1016/j.biopsycho.2014.09.010.

Mackey, A.P., Singley, A.T.M., Bunge, S.A., 2013. Intensive reasoning training alters pattern of brain connectivity at rest. J. Neurosci. 13, 4796–4803. doi:10.1523/JNEUROSCI.4141-12.2013.

Markovits, H., Brisson, J., de Chantal, P.-L., 2017. Logical reasoning versus information processing in the dual-strategy model of reasoning. J. Exp. Psychol.: Learn., Mem. Cognit. 43, 72–80. doi:10.1037/xlm0000291.

Markovits, H., Brunet, M.-L., Thompson, V., Brisson, J., 2013. Direct evidence for a dual process model of deductive inference. J. Exp. Psychol.: Learn., Mem. Cognit. 39, 1213–1222. doi:10.1037/a0030906.

Markovits, H., de Chantal, P.-L., Brisson, J., Gagnon-St-Pierre, E., 2019. The development of fast and slow inferential responding: evidence for a parallel development of rule-based and belief-based intuitions. Mem. Cognit. 47, 1188–1200. doi:10.3758/s13421-019-00927-3.

Mathôt, S., Schreij, D., Theeuwes, J., 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. Behav. Res. Methods 44, 314–324. doi:10.3758/s13428-011-0168-7.

Mazaika, P.K., Hoeft, F., Glover, G.H., Reiss, A.L., 2009. Methods and software for fMRI analysis of clinical subjects. Neuroimage 47, S58. doi:10.1016/S1053-8119(09)70238-1.

Monti, M.M., Osherson, D.N., 2012. Logic, language and the brain. Brain Res. 1428, 33–42. doi:10.1016/j.brainres.2011.05.061.

Noveck, I.A., Goel, V., Smith, K.W., 2004. The neural basis of conditional reasoning with arbitrary content. Cortex 40, 613–622. doi:10.1016/S0010-9452(08)70157-6.

Oaksford, M., 2015. Imaging deductive reasoning and the new paradigm. Front. Hum. Neurosci. 9, 101. doi:10.3389/fnhum.2015.00101.

Oaksford, M., Chater, N., 2007. Bayesian Rationality: the Probabilistic Approach to Human Reasoning. Oxford University Press, Oxford, England.

Oaksford, M., Chater, N., 2010. Conditional inference and constraint satisfaction: Reconciling mental models and the probabilistic approach. In: Oaksford, M., Chater, N. (Eds.), Cognition and Conditionals: Probability and Logic in Human Thinking. Oxford University Press, Oxford, England, pp. 309–333.

Oaksford, M., Chater, N., 2012. Dual processes, probabilities, and cognitive architecture. Mind Soc. 11, 15–26. doi:10.1007/s11299-011-0096-3.

Oaksford, M., Chater, N., 2020. New paradigms in the psychology of reasoning. Annu. Rev. Psychol. 71, 305–330. doi:10.1146/annurev-psych-010419-051132.

Oaksford, M., Chater, N., Larkin, J., 2000. Probabilities and polarity biases in conditional inference. J. Exp. Psychol.: Learn., Mem. Cognit. 26, 883–899. doi:10.1037/0278-7393.26.4.883.

Osherson, D., Perani, D., Cappa, S., Schnur, T., Grassi, F., Fazio, F., 1998. Distinct brain loci in deductive versus probabilistic reasoning. Neuropsychologia 36, 369–376. doi:10.1016/s0028-3932(97)00099-7.

Over, D.E., 2009. New paradigm psychology of reasoning. Think. Reason. 15, 431–438. doi:10.1080/13546780903266188.

Parsons, L.M., Osherson, D., 2001. New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. Cereb. Cortex 11, 954–965. doi:10.1093/cercor/11.10.954.

Pfeifer, N., 2013. The new psychology of reasoning: a mental probability logical perspective. Think. Reason. 19, 329–345. doi:10.1080/13546783.2013.838189.

Prado, J., Chadha, A., Booth, J.R., 2011. The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. J. Cogn. Neurosci. 23, 3483–3497. doi:10.1162/jocn_a_00063.

Prado, J., Noveck, I.A., 2007. Overcoming perceptual features in logical reasoning: a parametric functional magnetic resonance imaging study. J. Cogn. Neurosci. 19, 642–657. doi:10.1162/jocn.2007.19.4.642.

Prado, J., Van der Henst, J.-B., Noveck, I.A., 2010. Recomposing a fragmented literature: how conditional and relational arguments engage different neural systems for deductive reasoning. Neuroimage 51, 1213–1221. doi:10.1016/j.neuroimage.2010.03.026.

Ramsey, F.P., 1990. Truth and probability (original manuscript 1926). In: Mellor, D.H. (Ed.), Philosophical Papers. Cambridge University Press, Cambridge, England, pp. 52–94.

Ramsey, F.P., 1990. General propositions and causality (original manuscript, 1929). In: Mellor, D.H. (Ed.), Philosophical Papers. Cambridge University Press, Cambridge, England, pp. 145–163.

Reverberi, C., Cherubini, P., Frackowiak, R.S.J., Caltagirone, C., Paulesu, E., Macaluso, E., 2010. Conditional and syllogistic deductive tasks dissociate functionally during premise integration. Hum. Brain Mapp. 31, 1430–1445. doi:10.1002/hbm.20947.

Reverberi, C., Görgen, K., Haynes, J.-D., 2012. Distributed representations of rule identity and rule order in human frontal cortex and striatum. J. Neurosci. 32, 17420–17430. doi:10.1523/JNEUROSCI.1987-12.2012.

Rips, L.J., 1994. The Psychology of Proof. MIT Press, Cambridge, MA.

Rips, L.J., 2001. Two kinds of reasoning. Psychol. Sci. 12, 129–134. doi:10.1111/1467-9280.00322.

Rodriguez-Moreno, D., Hirsch, J., 2009. The dynamics of deductive reasoning: An fMRI investigation. Neuropsycholgia 47, 949–961. doi:10.1016/j.neuropsychologia.2008.08.030.

Rotello, C.M., Heit, E., 2009. Modeling the effects of argument length and validity on inductive and deductive reasoning. J. Exp. Psychol.: Learn., Mem. Cognit. 35, 1317–1330. doi:10.1037/a0016648.

Schwartz, F., Epinat-Duclos, J., Léone, J., Prado, J., 2017. The neural development of conditional reasoning in children: Different mechanisms for assessing the logical validity and likelihood of conclusions. Neuroimage 163, 264–275. doi:10.1016/j.neuroimage.2017.09.029.

Singmann, H., Klauer, K. C., 2011. Deductive and inductive conditional inferences: Two modes of reasoning. Thinking & Reasoning 17, 247–281.

Störring, G., 1908. Experimentelle Untersuchungen über einfache Schlussprozesse. Arch. Gesamte Psychol. 11, 1–27.

Svoboda, E., McKinnon, M.C., Levine, B., 2006. The functional neuroanatomy of autobiographical memory: a meta-analysis. Neuropsychologia 44, 2189–2208. doi:10.1016/j.neuropsychologia.2006.05.023.

Tsujii, T., Masuda, S., Akiyama, T., Watanabe, S., 2010. The role of inferior frontal cortex in belief-bias reasoning: an rTMS study. Neuropsychologia 48, 2005–2008. doi:10.1016/j.neuropsychologia.2010.03.021.

Vadeboncoeur, I., Markovits, H., 1999. The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task. Thinking & Reasoning 5, 97–113.

Verschueren, N., Schaeken, W., d'Ydewalle, G., 2005. A dual-process specification of causal conditional reasoning. Think. Reason. 11, 239–278. doi:10.1080/13546780442000178.

Weidenfeld, A., Oberauer, K., Hörnig, R., 2005. Causal and noncausal conditionals: An integrated model of interpretation and reasoning. Q. J. Exp. Psychol. 58A, 1479–1513. doi:10.1080/02724980443000719.

Wertheim, J., Ragni, M., 2018. The neural correlates of relational reasoning: A meta-analysis of 47 functional magnetic resonance studies. J. Cogn. Neurosci. 30, 1734–1748. doi:10.1162/jocn_a_01311.

Wertheim, J., Ragni, M., 2020. The neurocognitive correlates of human reasoning: A meta-analysis of conditional and syllogistic inferences. J. Cogn. Neurosci. 32, 1061–1078. doi:10.1162/jocn_a_01531.

Wilkins, M.C., 1928. The effect of changed material on the ability to do formal syllogistic reasoning. Arch. Psychol. 16 (102).

Woodworth, R.S., Sells, S.B., 1935. An atmosphere effect in syllogistic reasoning. J. Exp. Psychol. 18, 451–460. doi:10.1037/h0060520.

Yang, Y., Johnson-Laird, P.N., 2000. Illusions in quantified reasoning: How to make the impossible seem possible, and vice versa. Mem. Cognit. 28, 452–465. doi:10.3758/BF03198560.