



Reasoning strategies and semantic memory effects in deductive reasoning

Janie Brisson^{1,2} · Henry Markovits¹

Published online: 27 February 2020
© The Psychonomic Society, Inc. 2020

Abstract

Growing evidence supports the dual-strategy model, which suggests that reasoners have access to both a statistical and a counterexample reasoning strategy. In this paper, we explore further the processes underlying strategy use. We report three studies, the aim of which was to clarify the relation between this model and two forms of everyday reasoning. One of the most robust effects found with conditional reasoning with meaningful premises is the effect of alternative antecedents on the endorsement of AC and DA inferences. In a first study, we presented participants with conditional reasoning problems having more or fewer accessible alternatives as well as our dual-strategy diagnostic questionnaire. As hypothesized, results showed that strategy use had an independent effect on the inferences made with the AC and DA forms, over and above the effect of the number of antecedents, but was not related to responding to the MP and the MT forms. In a second study, we found that this relation extended to reasoning from an incompatibility statement. Finally, a third study showed that this relationship did not hold with probabilistic rather than logical response instructions, suggesting that the way reasoners transform a probabilistic evaluation into a dichotomous judgment is a key determinant of strategy use.

Keywords Dual strategy · Deductive reasoning · Working memory

Introduction

Deductive reasoning involves making inferences of logical validity based on a given set of premises. Although in theory such reasoning should depend solely on logical form, a great deal of research has clearly shown that humans are strongly influenced by the specific content of the premises (e.g., Cummins, Lubart, Alksnis, & Rist, 1991; Markovits & Vachon, 1990; Thompson, 1994). The clearest forms of such effects have been found in conditional (if-then) reasoning. Such reasoning has been mostly investigated through four inferential modes. According to the logical definition of the conditional connective, also called the material conditional, two of these inferences are valid and two are invalid. The *Modus Ponens* inference (MP: If P then Q, P is true, therefore

Q is true) and the *Modus Tollens* inference (MT: If P then Q, Q is false, therefore P is false) are both valid and lead to necessary conclusions. By contrast, the Affirmation of the Consequent inference (AC: If P then Q, Q is true, therefore P is true), and the Denial of the Antecedent inference (DA: If P then Q, P is false, therefore Q is false) are both invalid since their putative conclusion doesn't follow necessarily from the premises.

One of the most robust effects related to content is the effect of information stored in semantic memory on conditional reasoning. There are two classes of such information that have been shown to affect conditional reasoning. Disabling conditions refer to exceptions to the rule, that is, potential situations that can allow P to be true without Q being true (Cummins, 1995; Cummins et al., 1991). For example, for the conditional relation “If a rock is thrown at a window, the window will break,” one possible disabling condition would be to have a window made of plexiglass. Studies have shown that when there are more potential disabling conditions readily available in memory, the rejection rate for both the MP and the MT inferences increase (Byrne, 1989, 1991; Cummins, 1995; Cummins et al., 1991; Thompson, 1994), although this effect can be countered by stronger instructions (De Neys, Schaeken, & d’Ydewalle, 2003). The second class of

✉ Janie Brisson
brisson.janie@gmail.com

¹ Department of Psychology, Université du Québec à Montréal, Montreal, Quebec, Canada

² LaPsyDE, Unité CNRS 3521, Université Paris Descartes, Labo A. Binet, 46, rue Saint Jacques, 75005 Paris, France

information, which will be the major focus of these studies, concerns potential alternatives to the antecedent, which refer to cases A, such that A implies Q is true. For example, for the conditional relation “If a rock is thrown at a window, the window will break,” the case “throwing a chair at a window” would be a possible alternative antecedent. As for disabling conditions, the relative number of alternative antecedents readily available in memory has a strong effect on responses to both the AC and DA inferences. In both cases, the “logical” response is that no conclusion can be drawn with certainty. However, the rate of uncertainty responses is a linear function of the number of available alternative antecedents (Cummins, 1995; Cummins et al., 1991; Markovits & Vachon, 1990; Thompson, 1994). Starting from the example above, this would mean that reasoners would tend to accept the AC inference for a premises such as “If a rock is thrown at a window, the window will break” (i.e., they would conclude that if a window is broken, then a rock was thrown) less often than with the premise “If a finger is cut, then it will bleed” (i.e., they would conclude that if a finger bleeds, then it has been cut). The reason is that the former premise allows for many alternative antecedents, like throwing a chair, a car accident, a tropical storm, etc., that are counterexamples to the putative conclusion, while the latter contains many fewer of such antecedents (a finger is crushed, etc.). There is clear evidence that these effects are due to on-line access of semantic memory during reasoning (see Markovits, 2014, for a review). In contrast to the effect of disabling conditions, variation in responding to the AC and DA inferences as a function of the number of alternative antecedents is very robust. To our knowledge, even strong logical instructions have a limited effect on the relation between such alternatives and the rate of rejection of the AC and DA inferences (Evans, Handley, Neilens, & Over, 2010). In addition to this, recent evidence shows that the relation between access to alternatives and responding to uncertain inferences goes beyond conditional reasoning (Brisson, Markovits, Robert, & Schaeken, 2018).

What these studies clearly indicate is that familiar content provides implicit information that is accessed and subsequently influences the way that people make inferences, for what are otherwise identical logical forms. The question is then how is this information used when making inferences. Two major theoretical approaches to reasoning provide different explanations of the underlying processes. Probabilistic theories generally claim that reasoning involves generation of a likelihood estimate of the putative conclusion that relies on reasoners’ stored knowledge and beliefs (e.g., Evans, Over, & Handley, 2005; Oaksford & Chater, 2007). Within this context, it can be assumed that for conditional premises, the presence of more alternative antecedents decreases the probability of the corresponding conclusion. However the translation between conclusion probability and judgments of validity is subsequently done, a higher conclusion probability will translate into

increased endorsement in the context of validity judgments. By contrast, mental model theories postulate that when people reason, they construct internal models of the premises that reflect the semantics of the given logical connective (Johnson-Laird, 2001; Johnson-Laird & Byrne, 2002). However, they also acknowledge the possibility that some elements of these models will require access to memory, in particular those that include potential alternative antecedents, which in the case of the AC and DA inferences are counterexamples to the putative conclusions (see Markovits & Barrouillet, 2002). In this case, the relative number of such antecedents available in memory will determine the probability that at least one such element will be retrieved and incorporated into the internal model. If a counterexample is found, then the putative conclusion will be rejected.

Although each theoretical framework suggests a unique explanation for the way that information is used in reasoning, they do suggest two modes of information processing related to premises. The first of these uses a rapid associative access to information available in semantic memory in order to generate a likelihood estimate of a putative conclusion. The second examines the same information using a working memory intensive process that focusses on potential counterexamples to a putative conclusion.

A recent dual strategy model of reasoning (based on Verschueren et al., 2005a, b) suggests that people have access to two different modes of processing information, which roughly correspond to each of these models (Markovits, Forgues, & Brunet, 2012). One mode has been referred to as a statistical strategy and involves using associative access to information stored in semantic memory that is related to premises (e.g., various alternative antecedents from a conditional premise) in order to generate a likelihood estimation of a putative conclusion, similar to processes postulated by probabilistic theories. The second mode, referred to as a counterexample strategy, involves having access to the same information in order to generate an explicit internal representation of the premises with a corresponding search for counterexamples, similar to the processes postulated by the mental model theory. Several studies have now established a strong empirical basis for this dual-strategy model (Markovits, Brunet, Thompson, & Brisson, 2013; Markovits, Brisson, & De Chantal, 2015; Markovits, Brisson, De Chantal, & Thompson, 2017; Markovits, Trémolière, & Blanchette, 2018; Markovits, Brisson, De Chantal, & Singmann, 2018). For example, when people are given very short processing times, they will preferentially employ a statistical strategy. In addition, there is evidence that these strategies are not directly related to “logical” reasoning, but are basically characterized by different approaches to the way that information is processed (Markovits, Brisson, & de Chantal, 2017).

One important open question is the extent to which reasoning strategies might impact the way that people reason with familiar

premises with many (e.g., “If a rock is thrown at a window, the window will break”) and few (e.g., “If a finger is cut, then it will bleed”) alternative antecedents. Both strategies require access to the same implicit information about potential alternative antecedents. However, there are multiple models of how this might impact the interaction between reasoning and strategy. The simplest model would be that the number of potential alternative antecedents will affect only the extent to which this form of information is accessed during reasoning, irrespective of the strategy used by a reasoner. If it were assumed that this information was treated similarly, this would suggest that individual differences in strategy use should not affect the relationship between the number of alternatives and reasoning.

However, there is good evidence that strategy does affect the way that people use statistical information when responding to the uncertain inferences. The clearest indication of this comes from studies that have used premises with explicit statistical information (Markovits et al., 2017). The clearest case is premises that are presented with a very small number of potential counterexamples, but for which the probability of the invited conclusion is nonetheless very high. In this case, results show that counterexample reasoners show a strong tendency to reject conclusions, while statistical reasoners tend to accept these same conclusions, based on the very same information. This suggests one of two possible models for reasoning with the uncertain inferences when potential alternative antecedents are only implicit. One such model would suggest that counterexample reasoners should reject all AC and DA conclusions at an equally high rate, since potential counterexamples are available for all such premises, while statistical reasoners modulate their estimate of the likelihood of conclusions as a direct function of a number of available alternatives. A more complex model would suggest a two-part process. Initially, reasoners must retrieve and activate potential alternative antecedents that are stored in memory. The probability of activating any such information would depend at least partly on the quantity of available alternatives. Thus, both statistical and counterexample reasoners would activate the network of alternatives at a higher rate for premises with greater relative numbers of alternative antecedents. This would produce an overall effect of alternatives for both forms of reasoning. However, counterexample reasoners would use any such information to reject the uncertain inferences, while statistical reasoners would use the same information to reject these inferences at a lower rate. Thus, we would expect an effect that would be a combination of the number of available alternatives and strategy use.

Given the clear evidence that activation and retrieval of implicit information are implicated in reasoning (Markovits, 2014) and the evidence that strategy influences the way that such information is processed (Markovits et al., 2017), we hypothesize that the combined model will best represent the interaction between accessibility of alternatives and strategy use.

Study 1

In this initial study, we examine the effects of a number of alternatives and strategy use in conditional reasoning with familiar premises. We specifically predict that (1) in line with many previous studies, rates of acceptance of the AC and the DA inferences will be higher for premises that allow for relatively few alternative antecedents and (2) rates of acceptance of the AC and the DA inferences will be generally higher for statistical reasoners than for counterexample reasoners. We also presented reasoners with the two certain inferences (MP, MT). In these cases, there is some relationship between implicit disabling information and acceptance level of these two inferences. We decided to reduce the variability associated with these two certain inferences by using straightforward logical instructions that have been shown to reduce the effect of disablers, and by using premises for which there were generally relatively few disablers available. This was done to allow dissociation between any general tendency towards more or less logical reasoning and the more specific forms of information processing postulated by the dual-strategy model.

Method

Participants A total of 104 students (36 men, 67 women; average age 23 years, 2 months; age range 17–68) participated in this study. All participants were native French speakers and volunteers and were recruited in colleges or universities in Montreal.

Materials Four paper and pencil booklets were prepared. On the first page of each booklet, participants were asked to give basic demographic information.

In the first booklet, participants were given two series of inferences, one of which presented conditional inferences with meaningful content with many available alternatives, while the other series presented the strategy assessment problems. The second booklet presented the same two series of inferences, except that the inferences with meaningful content had few available alternatives. In the two booklets, all premises had few disabling conditions. Premises were selected from prior studies, where their content had been pretested (Janveau-Brennan & Markovits, 1999; Quinn & Markovits, 1998; Markovits & Quinn, 2002).

We made one replication of each booklet where the order of strategy assessment and meaningful inferences was inverted.

Conditional reasoning task Conditional inferences with meaningful premises were presented in the following way. First, participants were given the following instructions (translated from the original French):

“In the following pages, we are going to show you some rules that you must *suppose to be true*. You have to assume that the rules are always true. For each rule, we are also going to show you some observations. Your task is to select the conclusion that follows logically from the rule and the given observation.”

On the top of each page, one major premise was presented. On the same page, four logical problems corresponding to the MP, MT, AC, and DA inferences were presented in a random order (which was variable for each major premise). In the first booklet, three major premises with many alternatives were presented. These were, in order:

“If a person is frightened, then the person will be sweaty;
If a person falls into a lake, then the person will be wet;
If a person reads a printed newspaper, then the person’s hands will be dirty.”

The second booklet presented three sets of conditional problems based on the premises with few alternatives. These were, in order:

“If a rail is misplaced, then the train will derail;
If a person cuts a finger, then the finger will bleed;
If the light increases, then the pupil will contract.”

For each inference, participants had to choose among three possible conclusions. The following is an example of an AC inference problem with many alternatives:

“Suppose that it is always true that:

If a person is frightened, then the person will be sweaty.

For each of the following observations, select the conclusion that follows logically from the rule and the given observation:

Peter is sweaty. One can conclude that:

1. Peter is frightened.
2. Peter is not frightened.
3. One cannot conclude whether or not Peter is frightened.”

In summary, Premise-type (*Many, Few*) was a between-subjects variable and Logical form was a within-subjects variable.

Strategy assessment task First, participants were given the following instructions (translated from the original French):

“Imagine that team of scientists are on an expedition on a recently discovered planet called Kronus. On the following pages, we will ask you to answer the question about phenomena that are particular to this planet. For each problem, you will be given a rule of the form if ... then that are true on Kronus according to the scientists. It is very important that you suppose that each rule that is presented is always true. You will then be given additional information and a conclusion that you must evaluate.”

The strategy assessment problems presented the set of 13 problems used by Markovits et al. (2012). Each problem described a causal conditional relation involving nonsense terms or relations that included frequency information concerning

the relative numbers of not-p.q and p.q cases out of 1,000 observations. Participants were then given AC inferences and were asked to indicate whether the conclusion could be logically drawn from the premises or not. The second problem set was identical to the first set, except that the content of the major premise was changed for each problem, while maintaining the same frequency information.

Of the 13 items, five had a relative frequency of alternative antecedents that was close to 10% (each individual item varied between 8% and 10%), five had a relative frequency that was close to 50% (each individual item varied between 48% and 50%), and three had a relative frequency of alternative antecedents that was presented as 0% (these last were presented in order to provide greater variability in problem types). The following is an example:

“A team of geologists on Kronus have discovered a variety of stone that is very interesting, called a TrolYTE. They affirm that on Kronus, if a TrolYTE is heated, then it will give off Philoben gas.

Of the 1,000 last times that they have observed TrolYtes, the geologists made the following observations:

910 times Philoben gas has been given off, and the TrolYTE was heated.

90 times Philoben gas has been given off, and the TrolYTE was not heated

From this information, Jean reasoned in the following manner:

The geologists have affirmed that: If a TrolYTE is heated, then it will give off Philoben gas.

Observation: A TrolYTE has given off Philoben gas.

Conclusion: The TrolYTE was heated.”

Procedure Each participant was randomly allocated one of the booklets, was told to take as much time as needed, and took part in the experiment individually.

Results and discussion

All statistical analyses were run using the IBM SPSS statistics software. We first analyzed performance on the strategy-assessment problem set. Forty-seven participants rejected all of the 10% alternative antecedent problems and all of the 50% alternative antecedent problems and were put into the *Counterexample* category. For 38 participants, acceptance rates on the 10% alternative antecedent problems were greater than those on the 50% alternative antecedent problems, so they were put into the *Statistical* category. A total of 19 participants were not grouped into one of these two categories, and these were eliminated from subsequent analyses.

We then encoded the conditional reasoning problems by conclusion acceptance (1 for acceptance and 0 for the “cannot conclude” response). We then performed a mixed model ANOVA on conclusion acceptance on the conditional

reasoning task, with Validity (Valid, Invalid) and Logical-form (Affirm, Deny) as repeated-measures and Strategy (Counterexample, Statistical), Premise-type (Few, Many) and Order (First, Second) as between-subject variables. All post hoc analyses were made using the Tukey procedure with $p = 0.05$. It should be mentioned that the dependent variables, that is, endorsements of logical forms, were count variables with four possible values (0, 1, 2, 3). Distributions with these types of variable are often skewed. The analysis of variance used here is generally robust, and has often been used in analyzing similar forms of reasoning; nonetheless, its potential limitations should be mentioned.

First, a main effect of Premise-type showed that participants endorsed inferences with few alternatives ($N = 51$, $M = 2.17$, $SE = 0.13$) more often than those with many alternatives ($N = 53$, $M = 1.46$, $SE = 0.12$), $F(1, 77) = 16.42$, $p < 0.001$, $\text{partial } \eta^2 = 0.176$. No significant effect of order was found, $F(1, 77) = 0.504$, $p = 0.48$. There was also a marginally significant Validity \times Premise-type interaction, $F(1, 77) = 3.64$, $p = 0.06$, $\text{partial } \eta^2 = 0.045$. Post hoc analyses revealed that invalid inferences with few alternatives were endorsed more often ($M = 1.73$, $SE = 0.14$) than those with many alternatives ($M = 0.81$, $SE = 0.16$), while the difference between the valid inferences with few ($M = 2.61$, $SE = 0.14$) and many ($M = 2.12$, $SE = 0.13$) alternatives was not significant. This pattern is in agreement with our manipulations since available alternative antecedents are relevant to the invalid inferences but not the valid ones. The classical content effects on conditional reasoning were thus replicated: the more available alternative antecedents, the less endorsement of invalid conclusions (see Table 1 for full results).

A main effect of Validity showed that participants endorsed the valid inferences ($M = 2.36$, $SE = 0.1$) more often than the invalid ones ($M = 1.27$, $SE = 0.11$), $F(1, 77) = 97.18$, $p < 0.001$, $\text{partial } \eta^2 = 0.555$. Moreover, we observed a significant Validity \times Logical-form interaction, $F(1, 77) = 19.47$, $p < 0.001$, $\text{partial } \eta^2 = 0.202$, showing that participants endorsed

the MP ($M = 2.55$, $SD = 0.87$) more often than the MT inference ($M = 2.21$, $SD = 1.09$). They also endorsed the DA ($M = 1.34$, $SD = 1.21$) more often than the AC inference ($M = 1.09$, $SD = 1.21$). We also observed a significant Validity \times Logical-form \times Premise-type interaction, $F(1, 77) = 9.26$, $p < 0.01$, $\text{partial } \eta^2 = 0.107$. This showed that the MT inference was endorsed less often with many alternatives than with few alternatives while this was not the case for the MP inference.

A Validity \times Strategy interaction, $F(1, 77) = 10.5$, $p < 0.01$, $\text{partial } \eta^2 = 0.118$, showed that counterexample reasoners endorsed invalid conclusions ($M = 0.99$, $SE = 0.14$) less often than statistical reasoners did ($M = 1.55$, $SE = 1.17$), with no difference observed on the valid conclusions. As predicted, strategy use had an independent effect on the inferences made with the AC and DA forms, over and above the effect of the number of antecedents, but was not related to responding to the MP and the MT forms (see Table 2 for full results).

No significant effect of Logical-form, $F(1, 77) < 1$, Strategy, $F(1, 77) = 1.34$, $p = 0.25$, nor Order, $F(1, 77) < 1$, were found. No significant Strategy \times Premise-type, $F(1, 77) < 1$, Validity \times Strategy \times Premise-type, $F(1, 77) = 1.97$, $p = 0.17$, Validity \times Logical-form \times Strategy, $F(1, 77) = 1.77$, $p = 0.19$, Logical-form \times Strategy, $F(1, 77) = 1.57$, $p = 0.21$, Logical-form \times Premise-type, $F(1, 77) < 1$, Logical-form \times Strategy \times Premise-type, $F(1, 77) < 1$, nor Validity \times Logical-form \times Strategy \times Premise-type, $F(1, 77) = 2.93$, $p = 0.09$, interaction was found.

Overall, these results are consistent with our hypothesis. They show the general replication of the effects of relative numbers of potential alternative antecedents on acceptance rates of the AC and the DA inferences. The one unusual result in this context is the fact that the MT inference was accepted less often when there were more alternative antecedents. Although not predicted, there are some studies that have shown a similar effect (Markovits & Doyon, 2004), which remains to be properly interpreted. They also show that counterexample reasoners reject the AC and the DA inferences at a higher rate than do statistical reasoners, while responding to

Table 1 Mean number of endorsed conclusions (out of three) for the four logical forms (MP, MT, AC, and DA) by Type (Many, Few) in Study 1

Premise-type	Logical form	Mean	Standard deviation
Many	MP	2.42	1.03
	MT	1.93	1.23
	AC	0.56	0.99
	DA	1.00	1.23
Few	MP	2.7	0.61
	MT	2.52	0.81
	AC	1.70	1.16
	DA	1.72	1.09

Table 2 Mean number of endorsed conclusions (out of three) for the four logical forms (MP, MT, AC, and DA) by Strategy (Counterexample, Statistical) in Study 1

Strategy	Logical form	Mean	Standard deviation
Counterexample	MP	2.62	0.80
	MT	2.23	1.05
	AC	0.72	1.08
	DA	1.15	1.22
Statistical	MP	2.47	0.95
	MT	2.18	1.16
	AC	1.55	1.22
	DA	1.58	1.18

the MP and the MT inferences shows no relationship to reasoning strategy.

Study 2

In the following study, we attempt to generalize our results to a different form of reasoning. Specifically, we examine the interaction between reasoning strategy and reasoning from an incompatibility statement (Brisson et al., 2018). From a logical point of view, when two propositions are incompatible, the truth of one necessarily leads to the falsity of the other, but critically, the falsity of one does not necessarily lead to the truth of the other. The inferences that we will use consist of two valid and two invalid logical forms. The valid ones are the Affirm First and the Affirm Second inference forms. The Affirm First inference involves reasoning with the premises “P is incompatible with Q, P is true” and leads to the logically correct conclusion, “Q is false” (e.g. “Being in Montreal is incompatible with being in Paris, John is in Montreal. Therefore, John is not in Paris”). The Affirm Second inference involves the premises “P is incompatible with Q, Q is true” and leads to the logically correct conclusion, “P is false.” The two invalid forms deny the first or the second proposition in the minor premise. We will refer to them as the Deny First and the Deny Second inferences. The Deny First inference involves reasoning with the premises “P is incompatible with Q, P is false.” Although these premises suggest the putative conclusion “Q is true,” the premises do not lead to a necessary conclusion. The Deny Second inference involves the premises “P is incompatible with Q, Q is false,” and also has no certain conclusion. For example, the premises “Being in Montreal is incompatible with being in Paris, John is not in Paris” suggests that “John is in Montreal,” but since John could be elsewhere, for example in Rome, one cannot be certain about the putative conclusion. The invited conclusions of the Affirm First and Affirm Second inferences should thus be endorsed as certain provided that the premises were true, while the invited conclusions of the Deny First and Deny Second inferences should be rejected. What then determines the endorsement of the invalid inferences? The key concept is the idea that for every pair of incompatible situations (P, Q), there exist others for which both P and Q are false. These third options are counterexamples to the invalid conclusions. The incompatibility inferences are relevant to everyday life since they are related to an informal fallacy, often found in politics and advertising, called the false dilemma fallacy. From a logical point of view, the false dilemma consists of falsely presenting a premise as an exclusive disjunction between two propositions. This can happen in two types of situation. On the one hand, when the fallacy discards “in-between” situations, the propositions are presented as being bound by an exclusive disjunction (either p or q), which allows only for one proposition to be true and the

other one false, while they are bound by an inclusive one (p or q, or both), which allows for both propositions to be true (Van Eemeren & Grootendorst, 2016). Being able to fully understand what is entailed by a situation of incompatibility may thus help to suspend judgment when a situation may seem like a dichotomy, but in reality is not (for a full presentation of incompatibility inferences, their relevance to everyday reasoning as well as the pretesting of the material, see Brisson et al., 2018).

For this study, we draw the hypothesis that we will observe a similar phenomenon to that in Study 1. That is, counterexample reasoners will use counterexamples more often to reject the invalid inferences (Deny First, Deny Second) than statistical reasoners will.

Method

Participants

A total of 212 students (74 men, 135 women (three participants failed to provide gender); average age 22 years, 8 months; age range 17–60) participated in this study. All participants were native French speakers and volunteers and were recruited in colleges or universities in Montreal.

Materials and procedure

Reasoning from an incompatibility task The incompatibility reasoning task was assessed using inference problems from Brisson et al. (study 2A, 2018). In this study, incompatibility statements were constructed using categories designed to produce variable numbers of third options. We first constructed incompatibility statements of the form “X is A is incompatible with X is B,” where A and B are base-level categories derived from the same parent category. For example, dogs and horses are both animals. The corresponding incompatibility would be that “For an animal, being a dog is incompatible with being a horse.” We constructed an initial set using parent categories that allowed many possible base-level categories, which we refer to as Broad statements. We then constructed a set of Reduced statements, where the Reduced parent was chosen to have relatively few base-level instances, for example polar animals. Finally, we constructed a set of Close to binary statements, for which the parent was an action category that allowed almost no third options, for example from the parent voting on a bill, which allows voting for or against or abstaining.

Each problem varied according to content and inference validity. We presented four sets of inferential problems, which consisted of the Affirm First, Affirm Second, Deny First, and Deny Second inferences. The major premises of the Broad category were, in order (translated from the original French):

“1. For a vegetable, being a broccoli is incompatible with being a pepper;

2. For an animal, being a dog is incompatible with being a horse;
3. For a person, being in Montreal is incompatible with being in Paris;
4. For a fruit, being a grape is incompatible with being a strawberry.”

Another set of inferential problems was based on the premises from the Reduced category. These were, in order (translated from the original French):

- “1. For a root vegetable, being a potato is incompatible with being a carrot;
2. For a polar animal, being a polar bear is incompatible with being a penguin;
3. For a fruit with pits, being a peach is incompatible with being a cherry;
4. For a dessert of the day in a restaurant, being the chocolate cake is incompatible with being the lemon pie.”

Finally, a third set of inferential problems was based on the premises from the Close to binary category. These were, in order (translated from the original French):

- “1. For a player at the “heads and tails” game, betting on tails is incompatible with betting on heads;
2. For a driver at a fork on the road, taking the right road is incompatible with taking the left road;
3. For a player at the “even and odd” game, betting on an even number is incompatible with betting on an odd number;
4. For a person who votes on a bill, voting for the bill is incompatible with voting against the bill.”

Participants were then given the following instructions (translated from the original French):

“In the following pages, you will be presented statements that you must *suppose to be true*. Your task is to select the conclusion that follows logically from the given statements.”

Before each problem set, a short context was presented into which the incompatibility statement was presented. On the same page, four logical problems were presented in the following order: Affirm First, Affirm Second, Deny First, Deny Second. For each problem, participants had to choose amongst three possible conclusions. The following is an example of such a context and an Affirm First inference problem in the Many Third options condition:

“In an encyclopedia, we learn about animals. An animal cannot be both a dog and a horse at the same time. In other words, being a dog is incompatible with being a horse.

Suppose it is always true that:

Being a dog is incompatible with being a horse.

An animal is a dog. One can conclude that:

1. This animal is a horse.
2. This animal is not a horse.
3. One cannot conclude if this animal is a horse or not.”

In summary, Premise-type (Broad, Reduced, Close to binary) was a between-subjects variable and logical form was a within-subjects variable.

Strategy assessment task. The strategy assessment problems were identical to those used in Study 1.

Results and discussion

Performances on the strategy assessment problem set were analyzed as they were in study 1. A total of 117 participants were grouped in the *counterexample* strategy and 65 in the *statistical* strategy. The remaining 30 participants were not grouped into the *statistical* nor the *counterexample* category and were thus eliminated from subsequent analyses.

We then encoded the incompatibility reasoning problems by conclusion acceptance (1 for acceptance and 0 for the “cannot conclude” response). We then performed a mixed model ANOVA on conclusion acceptance on the incompatibility reasoning task, with Validity (Valid, Invalid) and Logical-form (First, Second) as repeated-measures and Strategy (Counterexample, Statistical), Premise-type (Close to binary, Reduced, Broad) and Order (First, Second) as between-subject variables.

We observed a main effect of Premise-type, $F(2, 166) = 18.8$, $p < 0.001$ and a Validity \times Premise-type interaction, $F(2, 166) = 17.6$, $p < 0.001$, partial $\eta^2 = 0.175$. We found no significant effect of Order, $F(1, 166) < 1$. Post hoc comparisons used the Student-Newman-Keuls procedure with $p = 0.05$. For invalid inferences, the close to binary inferences were endorsed more often ($M = 2.64$, $SE = 0.2$) than the reduced ones ($M = 1.69$, $SE = 0.23$), which in turn were endorsed more often than the broad inferences ($M = 0.75$, $SE = 0.21$). Content effects with reasoning from an incompatibility statement (Brisson et al., 2018) were thus replicated. That is, the more available alternatives, the less endorsement of the invalid inferences (see Table 3 for full results).

Table 3 Mean number of endorsed conclusions (out of four) for the four logical forms (Affirm First, Affirm Second, Deny First, and Deny Second) by Type (Broad, Reduced, Close to binary) in Study 2

Premise-type	Logical form	Mean	Standard deviation
Broad	Affirm First	3.81	0.53
	Affirm Second	3.78	0.55
	Deny First	0.57	1.17
	Deny Second	0.54	1.20
Reduced	Affirm First	3.79	0.62
	Affirm Second	3.69	0.81
	Deny First	1.56	1.85
	Deny Second	1.60	1.86
Close to binary	Affirm First	3.63	0.78
	Affirm Second	3.70	0.74
	Deny First	2.60	1.61
	Deny Second	2.60	1.61

A main effect of Validity also showed that participants endorsed the valid inferences ($M = 3.7$, $SE = 0.05$) more often than the invalid ones ($M = 1.69$, $SE = 0.12$), $F(1, 166) = 204.01$, $p < 0.001$, $\text{partial } \eta^2 = 0.551$. A Validity \times Strategy interaction $F(2, 166) = 12.6$, $p < 0.001$ also showed that counterexample reasoners endorsed invalid conclusions ($M = 1.33$, $SE = 0.15$) less often than statistical reasoners did ($M = 2.03$, $SE = 0.2$) (see Table 4 for full results). These post hoc analyses used the Tukey procedure with $p = 0.05$.

A marginally significant main effect of Strategy suggested that overall, statistical reasoners endorsed conclusions ($M = 2.8$, $SE = 0.1$) more often than counterexample reasoners did ($M = 2.59$, $SE = 0.07$), $F(1, 166) = 3.02$, $p = 0.08$.

No significant effect of Logical-form, $F(1, 166) < 1$ and no significant Strategy \times Premise-type, $F(1, 166) < 1$, Validity \times Strategy \times Premise-type, $F(1, 166) < 1$, Logical-form \times Strategy, $F(1, 166) = 1.15$, $p = 0.29$, Logical-form \times Premise-type, $F(1, 166) = 1.19$, $p = 0.31$, Logical-form \times Strategy \times Premise-type, $F(1, 166) < 1$, nor Validity \times Logical-form \times Strategy \times Premise-type, $F(1, 166) = 1.82$, $p = 0.17$, interaction was found.

These results are consistent with our predictions and show the exact same pattern that was found in Study 1. First, they replicated previous results showing that responding to the Deny first and Deny second inferences reflected the relative number of third options. Second, strategy use had an independent effect on the Deny First and Deny Second inference forms, over and above the effect of third options, but was not related to responding to the Affirm First nor Affirm Second forms. More specifically, counterexample strategy users rejected the former inferences more often than statistical reasoners.

Study 3

Studies 1 and 2 are based on the idea that two types of processes are at play during deductive reasoning with meaningful

content: the retrieval of relevant information and the translation of this information into a validity judgment. The results of these studies support the idea that both counterexample and statistical strategies retrieve the same basic information, and that the dual-strategy model distinguishes individual differences in the way that this information is used to make deductive judgments. Counterexample reasoners use retrieved alternatives to a putative conclusion to reject this conclusion, while statistical reasoners use these alternatives to construct the likelihood of the conclusion being true, which results in a higher rate of conclusion endorsement.

This analysis relies on a hypothesized major difference between strategies in the way that information is translated into a dichotomous judgment of validity. Although our previous results are consistent with this idea, stronger supporting evidence would be provided by comparing strategy use with a form of judgment for which the translation between retrieved information and judgment should be the same irrespective of strategy. The clearest way of doing this is to ask for explicitly probabilistic judgments. In this case, both statistical and counterexample reasoners will produce likelihood judgments, thus eliminating a key difference between strategies when dichotomous deductive inferences are asked for.

In this third study, we replicated Study 1, with the key difference that we asked for the explicit probability of being true for the putative conclusion. Since the information required for making these judgments is the same set of potential alternatives as for the premises in Study 1, we first predict that the classical content effects will be replicated. Specifically, we predict that the probability of the invalid conclusions will be greater for premises having fewer alternatives. Critically, we also predict that all reasoners will use the same processes to make the likelihood estimates, irrespective of strategy use, which will result in equivalent judgments.

Method

Participants A total of 157 students (57 men, 98 women (two participants failed to provide sex); average age 23 years, 8 months; age range 17–69) participated in this study. All participants were native French speakers and volunteers and were recruited in colleges or universities in Montreal.

Materials and procedure

Conditional reasoning task Conditional reasoning problems were presented as in Study 1, the sole difference being that rather than using logical validity instructions, we asked participants to evaluate the conclusion likelihood on a scale of 0 to 100. One example of an AC inference in such a setting would be to ask participants to evaluate the likelihood of the conclusion “Peter is frightened” from the premises “If a

Table 4 Mean number of endorsed conclusions (out of four) for the four logical forms (Affirm First, Affirm Second, Deny First, and Deny Second) by Strategy (Counterexample, Statistical) in Study 2

Strategy	Logical form	Mean	Standard deviation
Counterexample	Affirm First	3.85	0.48
	Affirm Second	3.83	0.52
	Deny First	1.27	1.72
	Deny Second	1.24	1.73
Statistical	Affirm First	3.54	0.86
	Affirm Second	3.54	0.93
	Deny First	2.11	1.71
	Deny Second	2.16	1.72

person is frightened, then the person will be sweaty; Peter is sweaty,” by selecting a number between 0 and 100 on a scale.

Strategy assessment task The strategy assessment problems were identical to those used in Study 1 and Study 2.

Again, we made one booklet with many and another one with few alternatives. Each booklet was replicated with an inverted order of strategy assessment and meaningful inferences.

Results and discussion

Strategy use was determined as in Studies 1 and 2. A total of 86 participants were grouped in the *counterexample* strategy and 39 in the *statistical* strategy. The remaining 32 participants were not grouped into the *statistical* nor the *counterexample* category and were thus eliminated from subsequent analyses.

We then performed a mixed model ANOVA on conclusion likelihood estimates for the conditional reasoning task, with Validity (Valid, Invalid) and Logical-form (Affirm, Deny) as repeated-measures and Strategy (Counterexample, Statistical), Premise-type (Many, Few) and Order (First, Second) as between-subject variables. All post hoc analyses were made using the Tukey procedure with $p = 0.05$.

A main effect of Validity $F(1, 115) = 128.96$, $p < 0.001$ showed that valid conclusions were estimated as more likely ($M = 83.5$, $SE = 1.5$) than invalid conclusions were ($M = 61.8$, $SE = 1.7$). A Validity \times Logical Form interaction, $F(1, 115) = 60.23$, $p < 0.001$, showed that participants estimated the MP inference ($M = 91.5$, $SE = 12.5$) as more likely than MT ($M = 80.5$, $SD = 23.9$). They also estimated the DA inference ($M = 68.2$, $SD = 24.0$) as more likely than AC ($M = 60.2$, $SD = 26.7$). A Validity \times Logical Form \times Premise type, $F(1, 115) = 21.27$, $p < 0.001$, showed that the difference between likelihood estimates on invalid inferences was significant only for premises with many alternatives.

A main effect of Premise-type $F(1, 115) = 35.42$, $p < 0.001$ and a Validity \times Premise-type interaction, $F(1, 115) = 28.34$, $p < 0.001$ showed that the invalid inferences with few alternatives were estimated as more likely ($N = 78$, $M = 74.4$, $SE = 2.3$) than the invalid inferences with many alternatives ($N = 79$, $M = 49.1$, $SE = 2.5$) (see Table 5 for full results). The content effects observed in Study 1 were thus replicated. As available alternatives increased, likelihood estimates of the invalid conclusions dropped.

A main effect of Strategy showed that counterexample reasoners gave higher likelihood estimates ($M = 77.0$, $SE = 1.4$) than statistical reasoners did ($M = 68.3$, $SE = 2.1$), $F(1, 115) = 11.76$, $p < 0.01$ (see Table 6 for full results). Critically, and contrary to previous studies, we observed no significant Validity \times Strategy interaction, $F(1, 115) = 1.45$, $p = 0.23$.

Table 5 Mean likelihood estimates for the four logical forms (MP, MT, AC, and DA) by Type (Many, Few) in Study 3

Premise-type	Logical form	Mean	Standard deviation
Many	MP	91.56	12.86
	MT	76.81	27.42
	AC	41.06	21.48
	DA	59.10	24.37
Few	MP	91.38	12.27
	MT	83.79	19.87
	AC	77.26	18.03
	DA	76.44	20.66

There were no significant effects of Logical-form, $F(1, 115) = 3.2$, $p = 0.08$, and Order, $F(1, 115) = 0.821$, $p = 0.37$, and no significant interactions involving Strategy \times Premise-type, $F(1, 115) < 1$, Validity \times Strategy \times Premise-type, $F(1, 115) = 3.2$, $p = 0.08$, Logical-form \times Strategy, $F(1, 115) = 3.15$, $p = 0.08$, Logical-form \times Premise-type, $F(1, 115) = 1.17$, $p = 0.28$, Logical-form \times Strategy \times Premise-type, $F(1, 115) = 2.61$, $p = 0.11$, nor Validity \times Logical-form \times Strategy \times Premise-type, $F(1, 115) = 1.82$, $p = 0.17$.

In summary, Study 3 shows that the combined effect of strategy and alternatives observed with validity instructions did not hold with probabilistic instructions. As predicted, with probabilistic instructions, the translation between retrieved information and judgment was the same irrespective of strategy. This provides further support to the idea that the dual-strategy model identifies individual differences in the way information is used to make deductive judgments, but not probabilistic ones.

General discussion

The major goal of this paper was to investigate how the individual differences identified by the dual-strategy model are

Table 6 Mean likelihood estimates for the four logical forms (MP, MT, AC, and DA) by Strategy (Counterexample, Statistical) in Study 3

Strategy	Logical form	Mean	Standard deviation
Counterexample	MP	93.31	11.36
	MT	84.90	21.02
	AC	61.21	26.83
	DA	69.86	23.91
Statistical	MP	87.39	14.04
	MT	70.81	27.13
	AC	57.99	26.77
	DA	64.53	24.20

relevant to reasoning with meaningful premises. More specifically, we used premises with variable levels of implicit information about potential counterexamples to invalid inferences. We suggested a model for reasoning with these uncertain inferences, which contains two major parts. More specifically, we suggested that both statistical and counterexample reasoners would first activate and attempt to retrieve a network of alternatives stored in semantic memory. The probability of activating and retrieving a given network will vary according to the relative number of potential counterexamples. Statistical reasoners will directly use retrieved counterexamples in order to produce a likelihood estimate of a putative conclusion. In this case, it is possible to have a high likelihood with some counterexamples, which would tend to produce acceptance of the conclusion. By contrast, counterexample reasoners will process the same information in order to construct an internal model of the premises and the putative conclusion. The presence of a potential counterexample will thus be sufficient to reject the conclusion. This leads to the hypothesis that, all things being equal, counterexample reasoners should produce higher levels of rejections of putative premises.

Studies 1 and 2 supported our hypothesis. They showed that the translation of meaningful information into a judgment of validity is different according to the strategy. They supported the idea that counterexample reasoners use retrieved alternatives to a putative conclusion to reject this conclusion, while statistical reasoners use these alternatives to produce a likelihood estimate of the conclusion.

Importantly, Study 3 showed that the difference between the two strategies does not hold under instructions that require explicit likelihood judgements. This is again consistent with the idea that the major difference in the way that the two strategies process familiar reasoning lies in the process of translating implicit information into a dichotomous judgement of validity.

Although these results provide quite strong support for our basic model, it would be useful to investigate whether the same effects are observed with variations in disabling conditions (exception to the rule). This possibility should be investigated in further studies.

Overall, these results strongly support one major component of the dual-strategy model: the way reasoners transform a probabilistic evaluation into a dichotomous validity judgment is a key determinant of strategy use.

Moreover, Studies 1 and 2 have shown that the dual-strategy model identifies individual differences that carry over to reasoning with meaningful premises that are both similar (conditional in Study 1) and different (incompatibility in Study 2) to our strategy assessment tool. These results contribute to the generality of the processes identified by our model. Indeed, recent studies have shown that strategy use predicts performance in several areas beyond conditional reasoning like emotion processing (Markovits, Trémolière, &

Blanchette, 2018), spatial rotation (Markovits, 2019) and social reasoning (Gagnon-St-Pierre, Doucerain, & Markovits, submitted). Combined with these, our studies support the idea that the dual-strategy model represents an important individual difference in information processing.

Open practice statement Neither of the experiments reported in this article was formally preregistered. Neither the data nor the materials have been made available on a permanent third-party archive; requests for the data or materials can be sent via email to the lead author at brisson.janie@gmail.com.

References

- Brisson, J., Markovits, H., Robert, S., & Schaeken, W. (2018). Reasoning from an incompatibility: False dilemma fallacies and content effects. *Memory & Cognition*, 1–14.
- Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31(1), 61–83.
- Byrne, R. M. (1991). Can valid inferences be suppressed?. *Cognition*, 39(1), 71–78.
- Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23 (5), 646–658.
- Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, 19 (3), 274–282.
- De Neys, W., Schaeken, W., & D'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition*, 31(4), 581–595.
- Evans, J. S. B., Handley, S. J., Neilens, H., & Over, D. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *The Quarterly Journal of Experimental Psychology*, 63(5), 892–909.
- Evans, J. S. B., Over, D. E., & Handley, S. J. (2005). Suppositionals, extensionality, and conditionals: A critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychological Review*, 112, 1040–1052.
- Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology*, 35(4), 904.
- Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences*, 5(10), 434–442.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics and inference. *Psychological Review*, 109, 646–678.
- Markovits, H. (2014). Conditional reasoning and semantic memory retrieval. In A. Feeney, V. Thompson (Eds.). *Reasoning and memory*, Hove, UK: Psychology Press.
- Markovits, H. (2019). Reasoning strategy modulates gender differences in performance on a spatial rotation task. *Quarterly Journal of Experimental Psychology*, 72(12), 2870–2876.
- Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review*, 22(1), 5–36.
- Markovits, H., Brisson, J., & de Chantal, P. L. (2015). Additional evidence for a dual-strategy model of reasoning: Probabilistic reasoning is more invariant than reasoning about logical validity. *Memory & Cognition*, 43(8), 1208–1215.
- Markovits, H., Brisson, J., & de Chantal, P. L. (2017). Logical reasoning versus information processing in the dual-strategy model of

- reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(1), 72.
- Markovits, H., Brisson, J., de Chantal, P. L., & Singmann, H. (2018). Multiple layers of information processing in deductive reasoning: combining dual strategy and dual-source approaches to reasoning. *Journal of Cognitive Psychology*, 1–12.
- Markovits, H., Brisson, J., de Chantal, P. L., & Thompson, V. A. (2017). Interactions between inferential strategies and belief bias. *Memory & Cognition*, 45(7), 1182–1192.
- Markovits, H., Brunet, M. L., Thompson, V., & Brisson, J. (2013). Direct evidence for a dual process model of deductive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(4), 1213.
- Markovits, H., & Doyon, C. (2004). Information processing and reasoning with premises that are empirically false: Interference, working memory, and processing speed. *Memory & Cognition*, 32, 592–601.
- Markovits, H., Forgues, H. L., & Brunet, M. L. (2012). More evidence for a dual-process model of conditional reasoning. *Memory & Cognition*, 40(5), 736–747.
- Markovits, H., & Quinn, S. (2002). Efficiency of retrieval correlates with “logical” reasoning from causal conditional premises. *Memory & Cognition*, 30(5), 696–706.
- Markovits, H., Trémolière, B., & Blanchette, I. (2018). Reasoning strategies modulate gender differences in emotion processing. *Cognition*, 170, 76–82.
- Markovits, H. & Vachon, R. (1990). Conditional reasoning, representation, and level of abstraction. *Developmental Psychology*, 26 (6), 942.
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality*. Oxford, England: Oxford University Press.
- Quinn, S., & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: Strength of association as a predictive factor for content effects. *Cognition*, 68(3), B93–B101.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22 (6), 742–758.
- Van Eemeren, F. H., & Grootendorst, R. (2016). *Argumentation, communication, and fallacies: A pragma-dialectical perspective*. Routledge.
- Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005a). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning*, 11(3), 239–278.
- Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005b). Everyday conditional reasoning: A working memory-dependent tradeoff between counterexample and likelihood use. *Memory & Cognition*, 33(1), 107–119.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.