THINKING & REASONING
2020, VOL. 26, NO. 1, 1–30
https://doi.org/10.1080/13546783.2018.1552194

Routledge
Taylor & Francis Group

Check for updates

Advancing the specification of dual process models of higher cognition: a critical test of the hybrid model view

Bence Bago[a] and Wim De Neys[a,b]

[a]UMR 8240 LaPsyDÉ, Paris Descartes University, Sorbonne Paris Cité, Paris, France;
[b]UMR 8240, LaPsyDÉ, CNRS, Paris, France

ABSTRACT

Dual process models of higher cognition have become very influential in the cognitive sciences. The popular Default-Interventionist model has long favoured a serial view on the interaction between intuitive and deliberative processing (or System 1 and System 2). Recent work has led to an alternative hybrid model view in which people's intuitive reasoning performance is assumed to be determined by the absolute and relative strength of competing intuitions. In the present study, we tested unique new predictions to validate the hybrid model. We adopted a two-response paradigm with popular base-rate neglect problems in which base-rate information and a stereotypical description could cue conflicting responses. By manipulating the extremity of the base-rates in our problems we aimed to affect the strength of the "logical" intuition that is hypothesised to cue selection of the base-rate response. The two-response paradigm – in which people were required to give an initial response under time-pressure and cognitive load – allowed us to identify the presumed intuitively generated response. Consistent with the hybrid model predictions, we observed that experimentally reducing the strength of the logical intuition decreased the number of initial base-rate responses when solving problems in which base-rates and stereotypical information conflicted. Critically, reasoners who gave an initial stereotypical response were less likely to register the intrinsic conflict (as reflected in decreased confidence) in this case, whereas reasoners who gave an initial base-rate response registered more conflict. Implications and remaining challenges for dual process theorising are discussed.

ARTICLE HISTORY Received 7 August 2017Revised 8 November 2018Accepted 16 November 2018

KEYWORDS Dual process theory; default-interventionist model; hybrid model; conflict detection

Introduction

For centuries, human thinking has been portrayed as an interplay between intuitive and deliberate thought processes. This classic dichotomy is

CONTACT Bence Bago bencebagok@gmail.com
© 2019 Informa UK Limited, trading as Taylor & Francis Group

captured by so-called dual process models of higher cognition that have become very influential in modern day research on reasoning, judgment, and decision-making (Evans, 2010; Kahneman, 2011; Sloman, 1996; Stanovich, 2011). By now, dual process models have been used to explain numerous phenomena ranging from probabilistic or deductive reasoning biases (Kahneman, 2011), economic behavior (Alós-Ferrer & Strack, 2014), moral reasoning (Greene, 2013), cooperative behavior (Rand, Greene, & Nowak, 2012), and creativity (Barr, Pennycook, Stolz, & Fugelsang, 2015; Cassotti, Agogué, Camarda, Houdé, & Borst, 2016).

At the most basic level, a dual process model posits that there are two different types of thinking, often referred to as System 1 and System 2 processing. System 1 (also referred to as intuitive, heuristic, or Type 1 processing) operates fast and effortless whereas System 2 (also referred to as deliberate, analytic, or Type 2 processing) is believed to be slower and effortful[1]. There are different types of dual process models but arguably the dominant framework has been the serial or default-interventionist model that has been put forward by prominent scholars such as Daniel Kahneman (Kahneman, 2011) or Evans and Stanovich (2013).

At the core of the default-interventionist (DI) model lays a serial view on the interaction between System 1 and 2. The key idea is that when people are faced with a reasoning problem, they will typically rely on the fast System 1 to generate an answer. This is the default system. If needed, people can activate System 2 in a later phase to intervene and correct System 1 output (e.g., Evans, 2008; Evans & Stanovich, 2013). But this System 2 engagement only occurs after System 1 has been engaged and is not guaranteed. More generally, in the serial DI model, reasoners are conceived as cognitive misers who try to minimise cognitive effort (Kahneman, 2011). Since System 2 thinking is hard, people will often refrain from recruiting it and stick to the default System 1 response.

The serial DI model offers an appealing explanation for the widespread "bias" that has been observed in the reasoning and decision-making literature. To illustrate, consider the following problem (an adaptation of the famous base-rate problem, Tversky & Kahneman, 1974):

> There is an event with 1000 people. Jo is a randomly chosen participant who attended the event. We know that Jo is 23 years old and is finishing a

[1]Operation speed and effort are typical correlates of System 1 and 2 processing. The idea is that these features have often been associated with System 1 and 2 processing. But this does not necessarily need to be the case; the features do not necessarily need to align (e.g., a process might be effortless but slow, e.g., "incubation", Gilhooley, 2016[AQ]), and other features can be proposed to differentiate System 1 and 2 processing (e.g., "autonomy", Pennycook, 2017). See Evans and Stanovich (2013), and the debate between Melnikoff and Bargh (2018) and Pennycook, De Neys, Evans, Stanovich, and Thompson (2018), for an extensive discussion.

> degree in engineering. On Friday nights, Jo likes to go out and drink beer. We also know that 997 people attending the event were women. What is most likely: Is Jo a man or a woman?

In the problem, people get information about the composition of a sample and a short description of one participant (e.g., "Jo"). The problem is specifically constructed such that it cues a prepotent intuitive response based on heuristic, stereotypical associations prompted by the description (e.g., "Jo is an engineer and likes beer, so Jo is a man"). This is the intuitive *"heuristic"* default response that is believed to be generated by System 1. However, this response conflicts with the response that is cued by the base-rate information. Indeed, given that there are hardly any males in the sample (3 out of 1000) logically speaking it will be much more likely that a randomly drawn individual will be female. Although it might be more likely that Jo is an engineer on the basis of the description alone (e.g., in general, there might be more male than female beer-loving engineers), the extreme base-rates should push the scale to the "female" side. However, decades of empirical studies have established that the vast majority of participants opt for the heuristically cued intuitive response and seem to neglect elementary logico-mathematical principles in this and a range of related tasks (e.g., Kahneman, 2011).

Why do intelligent adults so often violate basic logico-mathematical principles? Default-Interventionist theorists have highlighted that the key problem is that taking these principles into account typically requires demanding System 2 computations (e.g., Kahneman, 2011; Stanovich & West, 2000). When the fast System 1 has provided us with a response, most reasoners will refrain from engaging the effortful System 2. Consequently, they will not detect that their answer conflicts with more logical considerations (Kahneman, 2011). Put differently, these biased System 1 reasoners do not detect that they are being biased. The few people who do engage System 2 will override the initially cued intuitive System 1 response after their System 2 deliberation is completed and manage to give the correct logico-mathematical response.

The above illustrates how the serial interaction view in the default-interventionist dual process view makes at least two critical assumptions or hypotheses about people's reasoning performance. First, biased reasoners who give the heuristic System 1 response that conflicts with logico-mathematical principles will not detect that their response conflicts with these principles (e.g., Kahneman, 2011). Second, deliberate System 2 processing is assumed to be essentially corrective in nature: Sound reasoning in the case of conflict implies correction of the default intuitive response (e.g., Evans & Stanovich, 2013). We can refer to these hypotheses as the "bias blind spot" and "corrective" assumption, respectively (De Neys, 2017).

To avoid confusion, we will clarify a number of points before advancing. First, we use the label "correct" or "logical" response as a shortcut to refer to "the response that has traditionally been considered as correct or normative according to standard logic or probability theory". The appropriateness of these traditional norms can be questioned (e.g., see Stanovich & West, 2000, for a discussion). Under this interpretation, the heuristic response should not be labeled as "incorrect" or "biased". Note that when we present our specific predictions and results in the context of base-rate problems, we will adopt the potentially more neutral labels "base-rate" and "stereotypical" response to minimise misinterpretation. Nevertheless, it will be clear that readers should refrain from a blind literal reading of the labels.

In the same vein, one should refrain from equating System 2 processing and normative correctness. No dual process theorist has ever claimed that System 1 is always wrong and System 2 is always right. For example, it is crisp clear that adults can readily compute the answer to the problem "How much is 5 + 5?" without any deliberation. At the same time, System 2 does not universally produce the normative response. There can be situations in which too much deliberation will lead people astray (e.g., Reyna, 2004). Hence, the normative correctness of a response cannot be a defining feature of System 1 and 2. It is simply the case that the two features (i.e., whether a response has been generated by System 1 or 2 and whether it is normatively correct or not according to traditional standards) are often correlated in the type of problems we typically study in the reasoning and heuristics and biases field (Evans, 2010). For example, it has been the demanding System 2 nature of correct responding in problems such as the base rate neglect task, conjunction fallacy, belief bias syllogisms, and many others that has been used to account for the established correlation between "correct" responding and individual differences in cognitive capacity (e.g., Evans & Stanovich, 2013; Stanovich & West, 2000). When we talk about the "corrective assumption" or "bias blind spot" we talk about these reasoning problems where logic/probability based-responding has traditionally been considered to result from System 2 deliberation. But clearly, this does not imply that one can or should generally equate System 2 processing and normative correctness.

Finally, readers need to keep in mind that the claim about the serial nature of the System 1 and 2 interaction in the default-interventionist model concerns the postulated processing architecture during the core reasoning process. As De Neys (2017) put it, literally speaking, one might argue that a response to a reasoning problem can never be purely intuitive. That is, before System 1 can cue an intuitive response one will need to read or listen to the problem premises, for example. Such reading and comprehension processes may require deliberation and draw on the very same resources that System 2 requires. Consequently, one can argue that every reasoning process starts with initial System 2 activation. Likewise, one might argue that every reasoning process also ends with System 2 activation. That is, once reasoners have computed a response to a problem, they will need to verbalise or type down their answer. This answer production may also require System 2. In this sense, it can be said that even the serial default-interventionist model assumes that System 2 is always "on". But the idea here is that System 2 is in a "low-effort" mode in which it simply accepts the suggestions made by System 1 without checking them (Kahneman, 2011). Hence, it does not engage in any proper deliberation so its core function is not activated. In sum, it is useful to bear in mind that the serial default-interventionist claim concerns the processing during the actual "reasoning" stage and not the initial encoding of the preambles or the ultimate overt response production (De Neys, 2017).

The default-interventionist model and the corresponding bias blind spot and corrective assumptions have had far-reaching impact on theorising in the various fields that have adopted dual process models and, more generally, our view of human rationality (e.g., Gürçay & Baron, 2017; Stanovich & West, 2000). However, in recent years direct experimental testing of the core assumptions has pointed to fundamental issues. Pace the "bias blind spot" hypothesis, a range of studies have established that often biased reasoners do show bias sensitivity (e.g., Bonner & Newell, 2010; De Neys & Glumicic, 2008; Gangemi, Bourgeois-Gironde, & Mancini, 2015; Pennycook, Trippas, Handley, & Thompson, 2014; Stupple, Ball, Evans, & Kamal-Smith, 2011; but see also Aczel, Szollosi, & Bago, 2016; Mata, Ferreira, Voss, & Kollei, 2017; Travers, Rolison, & Feeney, 2016). In these studies, participants are presented with both traditional reasoning problems in which a cued heuristic response conflicts with a logical principle and control no-conflict problems. Small content transformations in the control versions ensure that the intrinsic conflict in the traditional version is removed. For example, a no-conflict problem of the above base-rate problem would simply switch the base-rates around (e.g., "There are 997 males and 3 females in the sample"). Everything else stays the same. Hence, in the control case both the description and base-rates cue the same response (i.e., "Jo is a man"). We can test people's bias or conflict sensitivity by measuring how they process these different versions. If biased reasoners are blind heuristic thinkers who do not take logical principles into account, then the fact that they conflict or not with the cued heuristic response should not impact their reasoning. However, the available evidence indicates that biased reasoners often do register conflict. For example, biased reasoners show increased response doubt – as reflected in lower confidence and slightly longer decision latencies, when they give a biased answer on the conflict problems (e.g., De Neys, Rossi, & Houdé, 2013; Gangemi et al., 2015; Pennycook, Trippas, et al., 2014; Stupple et al., 2011). They also show increased activation of brain areas that are supposed to mediate conflict and error detection (i.e., the Anterior Cingulate Cortex, e.g., De Neys, Vartanian, & Goel, 2008; Simon, Lubin, Houdé, & De Neys, 2015).

Critically, the bias or conflict sensitivity is also observed under severe time-pressure and cognitive load (Bago & De Neys, 2017a; Franssens & De Neys, 2009; Johnson, Tubau, & De Neys, 2016; Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2014; Thompson & Johnson, 2014). These time-pressure and load manipulations are used to experimentally "knock-out" System 2 deliberation. Since System 2 processing is time and cognitive resource demanding we can minimise its impact by restricting participants' response time or burdening their cognitive resources with a demanding concurrent task. This allows us to determine whether a certain effect is driven by System 1 or System 2. In sum, in direct contrast with the bias blind spot hypothesis, available evidence indicates that biased reasoners not only show sensitivity to logic/heuristic conflict, they do so intuitively on the basis of mere System 1 processing.

In addition, the corrective DI assumption is also being questioned. Recall that in the DI framework correct logico-mathematical responses in case of conflict are assumed to result from a correction of the heuristic System 1 response after System 2 deliberation (e.g., Evans & Stanovich, 2013; Kahneman, 2011). However, evidence is amassing that correct responses in these cases are also generated intuitively (e.g., Bago & De Neys, 2017a, 2018; Newman, Gibb, & Thompson, 2017). Most direct evidence for this claim comes from studies that adopt a two-response paradigm (Thompson, Turner, & Pennycook, 2011). In this paradigm, participants are asked to immediately respond with the first intuitive answer that comes to mind. Afterwards, they are allowed to take all the time they want to reflect on the problem and generate a final response. To make sure that the initial response is generated intuitively on the basis of System 1 processing, it has to be generated under stringent time-pressure and/or cognitive load (Bago & De Neys, 2017a; Newman et al., 2017). This procedure allows us to examine the time-course of response generation and establish empirically which response is generated by System 1. Studies that adopted this approach clearly indicate that many reasoners who give a correct final response (i.e., after System 2 deliberation was allowed) already managed to give this response in the initial response stage in which they had to reason intuitively. Hence, pace the corrective DI assumption, correct responders do not necessarily need to deliberate to correct a faulty intuition, their intuitive System 1 response is already correct.

In sum, we believe that evidence is amassing against the core predictions of the serial DI model (but see also Evans, 2017 – for an alternative view). Note that traditional competitors of the serial model do not fare any better in this respect. For example, the parallel model (Epstein, 1994; Sloman, 1996) posits that System 1 and System 2 are always engaged simultaneously from the start of the reasoning process. In theory, this model can account for biased reasoners' conflict sensitivity. However, just like the serial model it still assumes that cueing of the logical answer relies on System 2 deliberation. As we mentioned, evidence suggests that this can be done on to basis of mere System 1 processing. Therefore, a number of scholars (e.g., Bago & De Neys, 2017a; Ball, Thompson, & Stupple, 2017; Banks, 2017; Białek & De Neys, 2017; De Neys, 2012; Pennycook, 2017; Pennycook, Fugelsang, & Koehler, 2015; Thompson & Newman, 2017; Trippas & Handley, 2017) have recently called for a new dual process view which we can refer to as a "hybrid" model[2] (De Neys, 2017).

At the most general level, what sets the hybrid model view apart is that it entails that the response that is traditionally considered to be computed by System 2 can also be cued by System 1. Hence, System 1 is assumed to generate (at least) two different types of intuitive responses. For example, in the case of a classic reasoning task one of these is the traditional "heuristic" intuitive response that is based on semantic and other associations (e.g., the response cued by the stereotypical description in the base-rate problem). This is the exact same intuitive response that is also assumed to be cued by the serial (and parallel) model. The critical second response is what we can refer to as a "logical" intuitive response which is based on elementary knowledge of basic logical and probabilistic principles (e.g., the role of base-rates). The underlying idea here is that even biased reasoners implicitly grasp elementary logical and probabilistic principles and activate this knowledge automatically when faced with a reasoning task. This intuitive logical knowledge allows one to detect that the heuristic intuition is questionable in case of conflict without a need to engage in demanding System 2 computations.

Clearly, if people have indeed logical intuitions such as the hybrid model entails, one might wonder why they still predominantly opt for the heuristic response? A key point is that the different intuitions can vary in strength or activation level (De Neys, 2012; Pennycook et al., 2015; Trippas & Handley, 2017). Typically, the heuristic intuition will be stronger (i.e., have a higher activation level) than the logical one. The presence of a logical intuitive response allows reasoners to detect conflict, but it does not suffice for the

[2]We use the "hybrid" model label to refer to core features that seem to be shared – under our interpretation – by the recent theoretical proposals of these various authors. It should be clear that this does not imply that these proposals are completely similar. We are talking about a general family resemblance rather than full correspondence and focus on commonalities rather than the differences.

logical response to be selected as overt answer. In most cases, the heuristic intuition will dominate, and the modal reasoner will still be biased. But critically there can be individual variance in this respect. For some reasoners, the logical intuition might be so weak that they even fail to detect conflict (Pennycook et al., 2015). For others, the logical intuition can be stronger than the heuristic one (Bago & De Neys, 2017a; Pennycook et al., 2015). Consequently, these latter individuals will also manage to give the correct answer as their initial response without any further System 2 engagement.

Taken together, these ideas result in a model in which one's intuitive reasoning performance is determined by the absolute and relative strength of different intuitions (Bago & De Neys, 2017a; Pennycook, 2017; Pennycook et al., 2015). Whatever intuition has the highest absolute strength level gets selected as initial response. The relative difference determines the level or likelihood of experienced conflict. The more equal the activation strengths (i.e., the smaller the relative difference), the more pronounced the conflict experience will be. For example, an individual with a very strong heuristic intuition and a weak logical intuition should be less likely to detect conflict than an individual with a logical and heuristic intuition that are equally strong.

Pennycook et al. (2015) were the first to point to the critical interplay of the absolute and relative strength (or "speed" as they referred to it) difference of competing intuitions. In their three-stage model of analytic engagement they posited that "the probability of conflict detection is dependent on the relative speed at which the competing initial responses come to mind (Pennycook et al., 2015, p. 61)". Bago and De Neys (2017a, 2017b) similarly pointed out how the absolute and relative strength of different intuitions can determine our reasoning performance . It is this feature that we see as central to what we refer to as the hybrid model and propose to test in this study.

There is little doubt that the hybrid model captures the recent empirical conflict detection and correct intuitive response generation findings that the standard DI (or parallel) model struggles to account for. However, in and by itself this is not surprising. In a sense one might argue that the hybrid model is a post hoc postulation. It was specifically designed to account for the observed empirical findings. It did not predict these findings a priori (although see also Pennycook et al., 2015, for some initial tests). This is an important difference with the DI model. The DI model made clear and testable predictions (e.g., the bias blind spot and corrective assumptions) that allowed us to test and validate or falsify the model. In order to advance the development of the hybrid model, we need to derive such a priori hybrid model predictions and test them empirically. In the present paper, we present a study that focuses on this issue.

One way to test and validate the hybrid model is by experimentally manipulating the strength of the logical intuition. One can achieve this, for example, by manipulating the extremity of the base-rates (e.g., Pennycook, Fugelsang, & Koehler, 2012; Pennycook et al., 2015). Extreme base-rates (e.g., 997 women and 3 men) present a stronger cue with respect to the importance of taking the base-rates into account than more moderate base-rates (e.g., "700 women and 300 men"). Logically speaking, the more dominant the larger group is in size, the more likely that a randomly drawn individual will belong to it. Hence, by manipulating the extremeness of the base-rates, we should affect the strength of the logical intuition; it will become weaker as the base-rate probabilities become more moderate (Pennycook et al., 2015).

This leads to at least three testable predictions. A first hybrid model prediction is that if the logical intuition is made weaker, we should observe fewer initial base-rate responses. This prediction is based on the postulation that the absolute strength level determines the initial response selection. Whatever intuition dominates gets selected. Hence, all other things being equal if we make the logical intuition less strong it will be even more likely that the competing heuristic intuition will dominate. Consequently, intuitive base-rate responses should be less likely. Second, stereotypical responders should be less likely to detect conflict when the logical intuition is less strong. This prediction is based on the assumption that the relative strength difference determines the conflict detection likelihood. In case of a dominant heuristic intuition, making the logical intuition less strong will increase the relative difference between the two (i.e., the heuristic will dominate even more) which should decrease the likelihood and/or the level of experienced conflict.

Interestingly, at first sight, the initial studies in which Pennycook et al. (2012, 2015) introduced the base-rate extremity manipulation might seem to support these predictions. In contrast with the extreme base-rate condition – in which the logical intuition strength should be maximal – the moderate-base-rate condition gave rise to fewer base-rate responses and less or no conflict detection effects. However, note that Pennycook et al. used a traditional "one-response" paradigm in which participants were allotted all the time they needed to deliberate and reflect on the problem. This implies that the results can be driven by System 2 processing. However, if the hybrid model is correct we should observe similar effects in the absence of any System 2 processing. That is, the claim is that people's intuitive reasoning performance is solely based on the absolute and relative intuitive strength differences within System 1. Hence, the reduced selection of the base-rate response and conflict detection effects should be observed in the absence of System 2 intervention.

To test the hypothesis we adopted Pennycook et al.'s base-rate extremity manipulation in the present study but combined it with a two-response design in which participants gave both an initial intuitive and final response after deliberation. In the initial response stage we imposed a challenging response deadline and a concurrent load task to guarantee that the findings could not be affected by System 2 processing. Key question is whether we will observe reduced base-rate responses and conflict detection at the initial, intuitive response stage.

Critically, the hybrid model makes a counter-intuitive but clear third prediction. In contrast with stereotypical responders, the few reasoners who still manage to give an intuitive base-rate response with moderate base-rates (i.e., when the logical intuition is made weaker) should show stronger conflict effects than with extreme base-rates (i.e., when the logical intuition is stronger). Why should this be the case? Figure 1 gives a pictorial illustration of the hybrid model assumptions. In the figure, we have plotted the strength of the different intuitions in imaginary activation "units". The bottom panel (1B) shows the modal case of stereotypical responders. This is the case we have focused on so far. The model assumes that stereotypical responders are "biased" precisely because their heuristic intuition is stronger (e.g., 4 units) than their logical intuition (e.g., 2 units). Now, imagine that our base-rate manipulation decreases the strength of a logical intuition with, say, 1 unit. This is illustrated at the right hand side of the figure. With moderate base-rates, a stereotypical responder's logical intuition strength will decrease (e.g., it will go from 2 units to 1 unit). Because of the logical strength reduction, the relative strength difference between the logical and heuristic intuition *increases* (e.g., it goes from a 2 to a 3 unit difference). Consequently, conflict detection becomes less likely for the stereotypical responders. But as the top panel of Figure 1 illustrates, we should expect the exact opposite effect for intuitive base-rate responders. The model assumes they opt for the base-rate response precisely because their logical intuition is stronger (e.g., 4 units) than their heuristic intuition (e.g., 2 units). With moderate base-rates, their logical intuition strength will decrease (e.g., it will go from 4 to 3 units). In this case, the experimental logical strength reduction will decrease the relative strength difference between the logical and heuristic intuition (e.g., it goes from a 3 to 1 unit difference). Consequently, since a smaller relative difference implies more conflict, reasoners who still opt for the base-rate response with moderate-base-rates should show a more pronounced conflict effect. These opposite effects of the logical strength manipulation on the intuitive conflict detection of reasoners who opt for the stereotypical and base-rate response should provide us with a strong test of the hybrid models' assumptions.

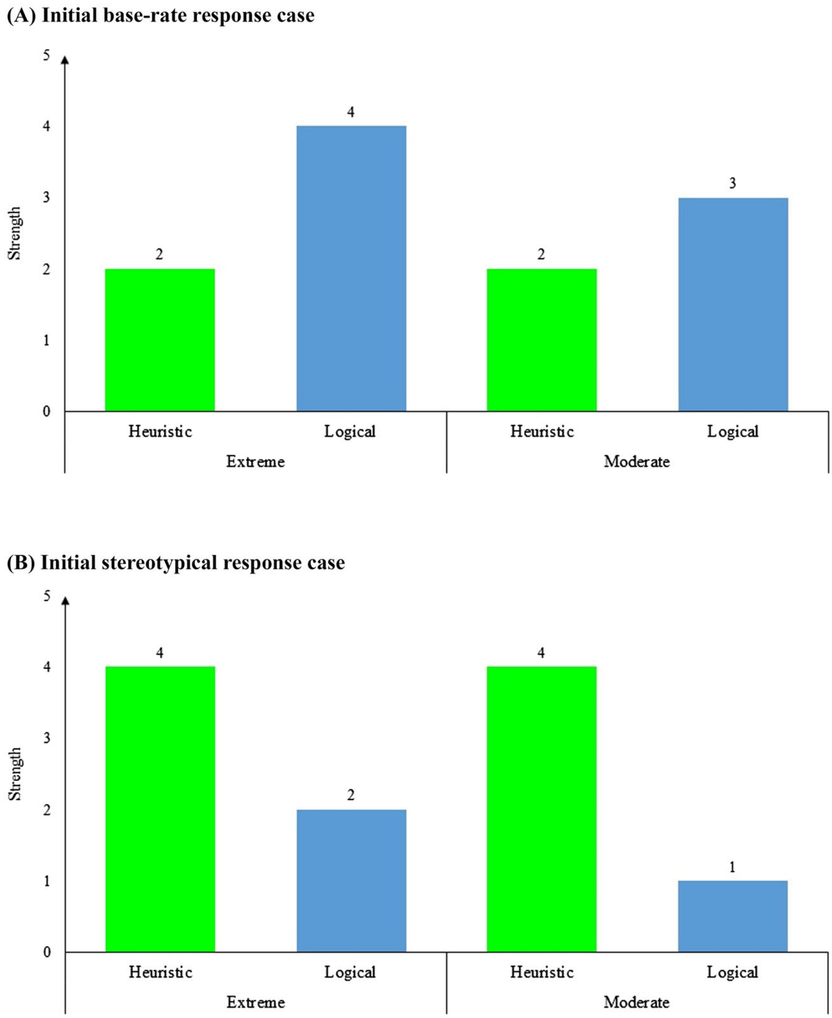

Figure 1. Illustration of the hybrid model predictions at the initial response stage for base-rate (A) and stereotypical (B) responders. The *y* axis represents the strength of the heuristic and logical intuition in imaginary strength units. Moderate base-rates are assumed to decrease the strength of the logical intuition by one unit.

Method

Participants

In total, 145 participants were tested (81 females, M = 40.7 years, SD =14.1 years). Participants were recruited online via the Crowdflower platform. Only North-American English speakers were allowed to participate. Participants were paid $0.25. A total of 40.6% of participants reported having high school as highest completed educational level, while 58% reported that they had a post-secondary educational degree (1.4% reported less than high school).

Material

Reasoning task. Participants solved eight base-rate problems. All problems were taken from Pennycook, Cheyne, Barr, Koehler, and Fugelsang (2014). Participants always received a description of the composition of a sample (e.g., "This study contained I.T engineers and professional boxers"), base-rate information (e.g., "There were 995 engineers and 5 professional boxers") and a description that was designed to cue a stereotypical association (e.g., "This person is strong"). Participants' task was to indicate to which group the person most likely belonged.

The problem presentation format we used in this research was based on Pennycook et al.'s (2014) rapid-response paradigm. In this format, the base-rates and descriptive information are presented serially and the amount of text that is presented on screen is minimised to minimise the influence of reading processes. Participants received three pieces of information in a given trial. First, the names of the two groups in the sample (e.g., "This study contains clowns and accountants") were presented. Second, participants were presented with stereotypical descriptive information (e.g., Person "L" is funny) as well. The descriptive information specified a neutral name ("Person L") and a single word personality trait (e.g., "strong" or "funny") that was designed to trigger the stereotypical association (based on extensive pretesting, see Pennycook et al., 2015). Finally, participants were also presented with the base-rate probabilities after the presentation of the stereotypes.

The following illustrates the full problem format:

This study contains clowns and accountants.
Person "L" is funny.
There are 995 clowns and 5 accountants.
Is Person "L" more likely to be:
o A clown
o An accountant

Half of the presented problems were conflict items and the other half were no-conflict items. In no-conflict items, the base-rate probabilities and the stereotypic descriptive information cued the same response (note that for convenience we will refer to this response as the "base-rate" response). In conflict items, the stereotypic information and the base-rate probabilities cued different responses (these are referred to as the stereotypical response and base-rate response, respectively). Two different item sets were used. The conflict items in one set were the no-conflict items in the other, and vice-versa. This was done by reversing the base-rates. Each of the two sets was used for half of the participants. This counterbalancing minimised the possibility that mere content or wording differences between conflict and no-conflict items could influence the results.

As in Pennycook et al. (2015), we used two kinds of base-rates (which were manipulated between-subjects): a moderate and extreme condition. As Pennycook et al. we also used three base-rate pairs within each condition: in the moderate condition they were 700/300, 710/290, 720/280, and in the extreme condition they were 997/3, 996/4, and 995/5. These slight manipulations of the base-rate pairs within each condition help to make the task less repetitive (De Neys & Glumicic, 2008). Only the base-rates were changed between the two conditions, everything else (stereotypes, name of the groups) remained constant. Participants were randomly allocated to the moderate or extreme base-rate treatment.

Each problem started with the presentation of a fixation cross for 1000 ms. After the fixation cross disappeared, the sentence which specified the two groups appeared for 2000 ms. Then the stereotypic information appeared, for another 2000 ms, while the first sentence remained on the screen. Finally, the base-rates appeared together with the question and two response alternatives. Note that we presented the base-rates and question together (rather than presenting the last information for 2000 ms first) to minimise the possibility that some participants would start solving the problem during the presentation of the last part of the problem. Once all the parts were presented, participants were able to select their answer by clicking on it. The position of the answer alternative that was cued by consideration of the base-rates/stereotype (i.e., first or second response option) was randomly determined for each item. The eight items were presented in random order.

Two-response format. The two-response task format was similar to the one introduced by Bago and De Neys (2017a). People were clearly instructed that we were interested in their first, initial response to the problem. Instructions stressed that it was important to give the initial response as fast as possible and that participants could afterwards take additional time to reflect on their answer. The literal instructions that were used stated the following:

> “Welcome to the experiment! Please read these instructions carefully!
> This experiment is composed of 8 questions and a couple of practice questions. It will take about 10 minutes to complete and it demands your full attention. You can only do this experiment once.
>
> In this task we’ll present you with a set of reasoning problems. We want to know what your initial, intuitive response to these problems is and how you

respond after you have thought about the problem for some more time. Hence, as soon as the problem is presented, we will ask you to enter your initial response. We want you to respond with the very first answer that comes to mind. You don't need to think about it. Just give the first answer that intuitively comes to mind as quickly as possible. Next, the problem will be presented again and you can take all the time you want to actively reflect on it. Once you have made up your mind you enter your final response. You will have as much time as you need to indicate your second response.

After you have entered your first and final answer we will also ask you to indicate your confidence in the correctness of your response. In sum, keep in mind that it is really crucial that you give your first, initial response as fast as possible. Afterwards, you can take as much time as you want to reflect on the problem and select your final response. You will receive $0.25 for completing this experiment. Please confirm below that you read these instructions carefully and then press the 'Next' button."

After the general instructions were presented the specific instructions for the base-rate task were presented:

"In a big research project a large number of studies were carried out where a psychologist made short personality descriptions of the participants. In every study there were participants from two population groups (e.g., carpenters and policemen). In each study one participant was drawn at random from the sample. You'll get to see one personality trait of this randomly chosen participant. You'll also get information about the composition of the population groups tested in the study in question. You'll be asked to indicate to which population group the participant most likely belongs. As we told you we are interested in your initial, intuitive response. First, we want you to respond with the very first answer that comes to mind. You don't need to think about it. Just give the first answer that intuitively comes to mind as quickly as possible. Next, the problem will be presented again and you can take all the time you want to actively reflect on it. Once you have made up your mind you enter your final response. After you made your choice and clicked on it, you will be automatically taken to the next page. After you have entered your first and final answer we will also ask you to indicate your confidence in the correctness of your response.

Press 'Next' if you are ready to start the practice session!"

After the task specific instructions, participants solved practice problems (specified under *"Procedure"*) to familiarise them with the task. Then they were able to start the experiment. For the first response, people were instructed to give a quick, intuitive response. After they clicked on the answer, they were asked to enter their confidence in their answer, on a scale from 0% to 100%, with the following question: "How confident are you in your answer? Please type a number from 0 (absolutely not confident) to 100 (absolutely confident)”. Next, they were presented with the problem again, and they were told that they could take as much time as they needed to give a final answer. As a last step, they were asked to give the confidence in their final answer.

The colour of the actual question and answer options were green during the first response, and they were blue during the second response phase, to visually remind participants which question they were answering at the moment. For this purpose, right under the question, a reminder sentence was placed: “Please indicate your very first, intuitive answer!” and “Please give your final answer.” respectively.

Response deadline. In order to minimise the possibility of System 2 engagement during the initial response, we used a strict response deadline (3000 milliseconds), based on a previous reading pre-test (see Bago & De Neys, 2017a). The deadline cut-off was based on the average time the pre-test participants needed to simply read the problems. 1000 ms before the deadline, the background turned yellow to alert the participants to the approaching deadline. If participants did not select an answer within 3000 ms they got feedback to remind them that they had not answered within the deadline and they were told to make sure to respond faster on subsequent trials. Obviously, there was no response deadline for the final response.

Cognitive load task. To further minimise the possibility of System 2 engagement during the initial response phase we also imposed a concurrent load task to burden participants’ cognitive resources (i.e., the dot memorisation task, see Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Here we also followed the procedure adopted by Bago and De Neys (2017a). The rationale behind the load manipulation is simple, System 2 processing is typically assumed to require executive cognitive resources, while System 1 processing does not (Evans & Stanovich, 2013). Consequently, if we burden participants’ executive resources while they are asked to solve reasoning problems, System 2 engagement is less likely. We opted for the dot memorisation task because it is well-established that it successfully burdens participant’s executive resources in a reasoning context (De Neys & Schaeken, 2007; De Neys & Verschueren, 2006; Franssens & De Neys, 2009; Johnson et al., 2016). Before each reasoning problem (and after the presentation of the fixation cross) participants were presented with a 3 x 3 grid, in which 4 dots were placed. Participants were instructed that it was critical to memorise the pattern even though it might be hard while solving the reasoning problem. After providing the initial response and the initial confidence rating, participants were shown four different matrixes and they had to choose the correct, to-be-memorised pattern. They received feedback as to whether they chose the correct or incorrect pattern. The load was only applied during the initial response stage and not during the subsequent final response stage in which participants were allowed to deliberate and recruit System 2.

Conflict detection measure. Previous research in the reasoning and cognitive control field established that the effect of conflict can be measured by the post-decision confidence level differences between conflict and no-conflict items (e.g., Botvinick, 2007; Johnson et al., 2016; Mevel et al., 2015; Pennycook et al., 2014; Stupple, Ball, & Ellis, 2013; Yeung & Summerfield, 2012). If people are being faced with two competing responses this should decrease their response confidence on the conflict items. Therefore, we will use this confidence difference as our primary index to measure the level of experienced conflict at the initial response stage; a higher confidence decrease is assumed to reflect a higher level of experienced conflict. Note that we refrained from using response latencies to measure conflict detection. Although this is a popular conflict measure in one-response paradigms, previous two-response studies established that it does not reliably track conflict detection effects reflected in confidence ratings at the initial response stage (Bago & De Neys, 2017a; Thompson & Johnson, 2014[3]). For completeness, one may also note that our assumption that the confidence decrease reflects the level of experienced conflict is questionable at the final response stage. After deliberate reflection, the initially experienced doubt can be mitigated (e.g., De Neys et al., 2013). However, our key interest lies in the initial, intuitive response stage in which System 2 deliberation is experimentally minimised.

Thus, confidence in the correctness of the response was recorded after the initial and the final response stage but our primary interest concerns the initial response stage. Note that participants were still under concurrent load while providing the initial confidence rating. This helps to guarantee that the confidence rating is not affected by post-decision System 2 processing.

Procedure

The experiment was run online. After the instructions, participants were presented with practice problems to familiarise them with the procedure. They first solved two practice (no-conflict) reasoning problem. After, they were presented with a practice dot matrix recall item (i.e., they were simply shown a dot pattern and after it disappeared they were asked to identify

[3]As Argued by Bago and De Neys (2017a) this might result from the specific design characteristics of the two-response paradigm. Forcing people to respond as fast as possible might prevent the slowing effect from showing up.

Table 1. Frequency of initial and final base-rate responses for conflict and no-conflict items with extreme and moderate base-rates.

		Base-rate extremity	
	Response	Moderate (%)	Extreme (%)
Conflict	Initial	16.4	29.7
	Final	23	41.6
No-conflict	Initial	90.9	93.4
	Final	90	93.7

the pattern from four presented options). As a last practice step, they were given two reasoning problems (the first of which was the initial reasoning problem) which they now had to solve under load. At the end of the experiment, standard demographic questions were recorded.

Exclusion criteria. All trials where participants did not manage to provide an initial response within the deadline were excluded from the analysis (9.1% of trials). We also excluded those trials where participants did not give the correct response to the dot memorisation task (17.4% of trials). These exclusion criteria help to guarantee that System 2 processing was maximally ruled out during the initial response stage. Altogether, we excluded 24.1% of trials and analyzed 881 trials (out of 1160).

Results

Frequency of base-rate response choices. Table 1 gives an overview of the frequency of the base-rate response choices in our study (i.e., percentage of trials on which the response cued by the base-rates was selected). Visual inspection of the table points to a number of expected trends. First, as in previous studies (e.g., De Neys & Glumicic, 2008; Pennycook et al., 2012, 2015) selection of the base-rate response is uniformly high on the no-conflict control trials (90% in moderate and 93% in extreme condition). This is not surprising. All dual process models predict that mere System 1 processing suffices to favor the base-rate response on these problems. Further in line with general expectations, we find that participants have a harder time on the traditional conflict problems: Reasoners typically opt for the stereotypical response here with a maximum base-rate response choice rate that does not exceed 42%. These effects replicate classic findings with the one-response paradigm. In line with previous two-response studies (e.g., Bago & De Neys, 2017a; Newman et al., 2017; Thompson et al., 2011) we also observe that there is a general trend towards slightly higher base-rate response choices at the final than at the initial response stage[4].

[4]As a side note, the current findings also replicate the two-response findings with respect to the non-corrective nature of System 2 processing (e.g., Bago & De Neys, 2017a). We observe that on the majority of trials on which a reasoner gives the base-rate response as their final answer, they

Turning to the effect of the base-rate extremity manipulation one can see that at the final response stage we replicate the findings of Pennycook et al. (2015): When reasoners had the time to deliberate, there are fewer base-rate responses with moderate (23%) than with extreme base-rates problems (41.6%). However, the key finding is that we observe a similar trend at the initial response stage: there are also fewer initial base-rate responses with moderate (16.4%) than with extreme (29.7%) base-rates.

To test these results statistically, we used mixed effect logistic regression models in which subjects were entered as random effect and response stage (initial or final) and base-rate extremity (moderate or extreme) as fixed factors. For the conflict problems, we found that response stage, χ^2 (3) = 20.49, $p < 0.0001$, $b = 1.09$, and base-rate extremity, χ^2 (4) = 9.27, $p < 0.0001$, $b = -2.32$, increased the model fit significantly but their interaction, χ^2 (5) = 0.998, $p = 0.32$, did not. These results confirm that the moderate (vs extreme) base-rate manipulation decreases the selection rate of base-rate responses at both response stages. As one may expect, none of these effects are observed on the no-conflict problems. Initial and final no-conflict base-rate selection rates are uniformly high. Neither the effect of response number, χ^2 (3) = 0.08, $p = 0.78$, nor base-rate extremity, χ^2 (4) = 1.06, $p = 0.30$, or their interaction, χ^2 (5) = 0.05, $p = 0.82$, was significant.

In sum, the general pattern of response choices is completely in line with previous studies. Key new finding is that we observe an effect of the base-rate extremity manipulation at the initial response stage. Moderate base-rates make intuitive base-rate response choices less likely. Hence, a manipulation that is assumed to decrease the strength of the logical intuition indeed leads to fewer initial responses that are assumed to be cued by this intuition. This supports the response choice prediction of the hybrid model.

Conflict detection findings. The hybrid model predicted that initial stereotypical responders will be less likely to experience conflict in the moderate base-rate condition, whereas initial base-rate responders should show the opposite effect and should experience more conflict when base-rates are moderate vs extreme. To test this hypothesis we contrasted the initial confidence levels for initial base-rate responses on no-conflict trials (which we will refer to as our "baseline") and the base-rate and stereotypical responses

already generated the base-rate response in the initial response stage (both with moderate and extreme base-rates, 61.5% and 64% of correct final trials, respectively). This provides further evidence against the corrective assumption: base-rate responders do not necessarily need to engage System 2 to correct their intuition, their intuitions already favors the base-rates. However, this does not imply that correction does not occur. Although most base-rate responders do not need to correct their intuitive response, some do. This is reflected, for example, in the higher overall base-rate response selection that we observed in the final vs initial response stage.

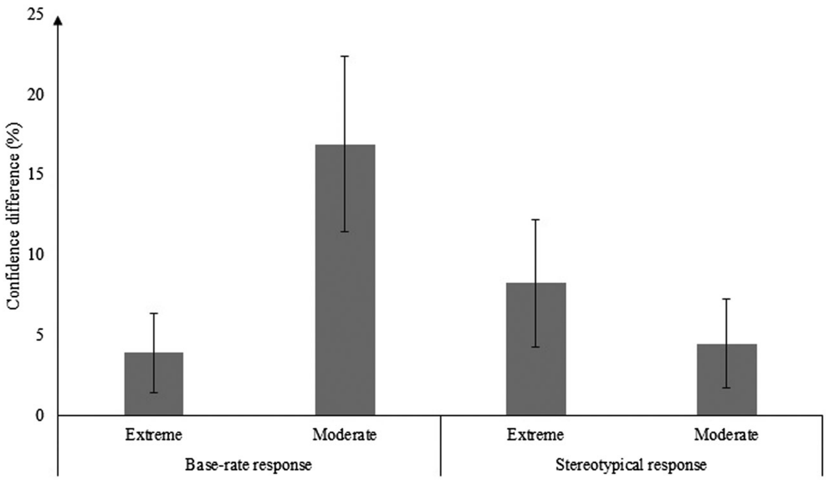

Figure 2. Initial confidence difference between the no-conflict baseline and conflict problems on which the base-rate or stereotypical response was chosen with extreme and moderate base-rates. A higher difference value reflects a more pronounced conflict detection. Error bars are standard errors of the difference between the means of the baseline and conflict cases.

Table 2. Overall confidence ratings (A) and confidence contrast difference between the no-conflict baseline and conflict problems (B) as a function of response stage, response choice and base-rate extremity.

		Base rate extremity	
Response	Response choice	Moderate	Extreme
A. Overall			
Initial	Base-rate no-conflict	80.7% (21.7)	88% (18.7)
	Other no-conflict	53.2% (41.1)	53.8% (35.2)
	Base-rate conflict	65.2% (31.9)	84.4% (16.3)
	Stereotypical conflict	77.6% (25.2)	79.8% (28.9)
Final	Base-rate no-conflict	83.7% (20.3)	91.6% (15.4)
	Other no-conflict	54.2% (38.4)	41.6% (31.9)
	Base-rate conflict	65.4% (32.1)	83.3% (24.5)
	Stereotypical conflict	79.8% (24.7)	83.9% (20.7)
B. Difference contrast (correct no-conflict baseline − conflict)			
Initial	Base-rate conflict	15.5% (5.5)	3.6% (2.5)
	Stereotypical conflict	3.1% (2.4)	8.2% (2.8)
Final	Base-rate conflict	18.3% (4.7)	8.3% (3.1)
	Stereotypical conflict	3.9% (2.3)	7.7% (2.6)

Standard deviations of the mean are in brackets.

on conflict trials. We discard the rare "other" no-conflict trials because these cannot be interpreted unequivocally (see Bago & De Neys, 2017a; De Neys et al., 2011; Pennycook et al., 2015).

Figure 2 shows the conflict detection effect findings (i.e., confidence contrast, no-conflict minus conflict trials confidence) at the initial response stage (see also Table 2 for complete overview). A higher difference value implies a larger confidence decrease when solving conflict items which should reflect a more pronounced conflict experience. Visual inspection indeed confirms the predictions of the hybrid model. For stereotypical responses we see a smaller confidence decrease or conflict effect with moderate base-rates than with extreme base-rates, whereas base-rate responses show the opposite trend. Hence, consistent with the hybrid model predictions, an experimental manipulation that decreases the strength of the logical intuition that is assumed to cue selection of the base-rate response tends to make stereotypical responders feel less conflicted and base-rate responders more conflicted.

We used mixed effect linear regression models to test the visual trends in the initial confidence data statistically[5]. We entered the random intercept of subjects in the models. As fixed factors, we entered a variable which we will refer to as "response group", base-rate extremity (moderate or extreme), and their interaction. The "response group" variable coded whether a given data point was a no-conflict trial on which the base-rate response was selected, a conflict trial on which the stereotypical response was selected, or conflict trial on which the base-rate response was selected. If base-rate extremity has opposite effects on stereotypical and base-rate responders' conflict experience we would expect a significant interaction between the response group and extremity factors. The analysis showed that the main effect of response group, χ^2 (5) = 27.12, $p < 0.0001$, improved model fit significantly, while the main effect of base-rate extremity did not, χ^2 (6) = 3.83, p = 0.0502. Critically, the interaction also improved fit further, χ^2 (8) = 11.43, p = 0.003. To follow-up on this interaction, we ran separate analyses for conflict trials on which the base-rate and stereotypical response was selected. Here we tested whether the simple interaction between the conflict factor (conflict or no-conflict) and base-rate extremity (moderate or extreme) was significant. This allows us to test whether the observed conflict effect *decrease* with moderate base-rates for stereotypical conflict responses and the observed conflict effect *increase* with moderate base-rates for base-rate responses were statistically significant. Results showed that this was indeed the case. Both for base-rate, b = 5.26, t (719.7) = 2.3, p = 0.021, and stereotypical responses, b = –8.06, t (748.2) = –2.1, p = 0.036, the interaction significantly improved model fit.

[5]Mixed effect models use the individual items as elementary unit of analysis. Contrary to a more classic ANOVA approach performance is not averaged across trials. Our use of the term "responder" literally refers to the performance on a single trial. Although people's two-response choices have been shown to be highly stable across trials (Bago & De Neys, 2017a, 2018), it is possible that a base-rate responder on trial x will be a stereotypical responder on trial y. Given that the strength of intuitions may vary across trial (e.g., stereotype x might cue a stronger response than stereotype y for individual z) such potential trial-by-trial variability is not problematic for the hybrid model (or – to our knowledge – any other dual process model).

Additional data. Our primary interest concerned the confidence conflict findings for the initial, intuitive responses. However, we also recorded confidence rating for the *final* response. For completeness, Table 2 (bottom panel) presents an overview of these findings. As the table indicates the pattern is similar to what we observed at the initial response stage. The differential impact of the extremity manipulation for base-rate and stereotypical responses is also present at the final response stage. However, the final response confidence findings should be interpreted with caution. First, the hybrid model we set out to test here made no clear prediction on what would happen at the final response stage. In addition, deliberate System 2 processing might mitigate the intuitively detected initial conflict. Indeed, especially for base-rate responses, final confidence cannot be considered a pure index of conflict detection per se (De Neys et al., 2011; Pennycook et al., 2015).

Nevertheless, for consistency, we used the exact same mixed effect regression models approach as with the initial confidence findings to test the final confidence trends statistically. As with the initial confidence data, there was a significant interaction between response group (base-rate response on no-conflict trial, base-rate response on conflict trial, or stereotypical response on conflict trial) and extremity factors (moderate vs extreme), χ^2 (8) = 22.75, $p < 0.0001$. In follow-up tests, we again tested the interaction between the conflict factor (conflict vs no conflict) and base rate extremity (extreme vs moderate) separately for base-rate and stereotypical conflict responses. Results showed that the interaction was significant both for the correct, $b = -12$, t (754.7) = –3.6, $p = 0.0003$, and incorrect, $b = 5.86$, t (725.2) = 2.5, $p = 0.012$, responses. Hence, the final confidence trends are consistent with the initial confidence ones but should be interpreted with caution.

With the same caveat in mind, we also present the descriptive conflict latency contrast data for initial and final responses. As we explained, response latencies are a popular conflict detection measure in one-response paradigms, but previous two-response paradigms indicated that they do not reliably track conflict detection effects reflected in confidence ratings at the initial response stage (Bago & De Neys, 2017a; Thompson & Johnson, 2014; but see also Bago & De Neys, 2018). As Table 3 indicates, this trend is also observed in the present study. At the initial response stage – where people are forced to give a response as fast as possible – the descriptive data do generally not point to longer processing times for conflict problems and do not track the confidence results. At the final response stage, the latency pattern also diverges from the confidence pattern. The same mixed model regression approach as with the confidence ratings indeed indicated that neither for the initial, χ^2 (8) = 4.5, $p = 0.08$, nor final

Table 3. Overall response times (A) and response time differences between the no-conflict baseline and conflict problems (B) as a function of response stage, response choice, and base-rate extremity. Standard deviations of the mean are in brackets.

		Base rate extremity	
Response	Response choice	Moderate	Extreme
A. Overall			
Initial	Base-rate no-conflict	1.43 s (1.4)	1.63 s (1.4)
	Other no-conflict	1.14 s (1.3)	1.61 s (1.5)
	Base-rate conflict	1.58 s (1.5)	1.96 s (1.3)
	Stereotypical conflict	1.41 s (1.4)	1.46 s (1.4)
Final	Base-rate no-conflict	2.67 s (1.6)	2.86 s (1.7)
	Other no-conflict	2.3 s (1.7)	2 s (1.5)
	Base-rate conflict	2.89 s (1.8)	2.36 s (1.9)
	Stereotypical conflict	2.65 s (1.7)	2.91 s (1.7)
B. Difference contrast (correct no-conflict baseline − conflict)			
Initial	Base-rate conflict	−0.15 s (0.26)	−0.33 s (0.2)
	Stereotypical conflict	0.02 s (0.14)	0.17 s (0.15)
Final	Base-rate conflict	−0.22 s (0.27)	−0.5 s (0.24)
	Stereotypical conflict	0.02 s (0.17)	−0.05 s (0.2)

Means were calculated on log-transformed data and were back-transformed prior to the subtraction.

response times, χ^2 (8) = 0.3, p = 0.86, the critical interaction term between response group and base rate extremity improved model fit.

However, one might note that at the descriptive level both the final latency and confidence contrast indexes do point to a less pronounced conflict detection effect for stereotypical responses at the final response stage (i.e., less pronounced slowing and response doubt) with moderate vs extreme base-rates. Although the exploratory ad hoc nature of these additional data analyses needs to be kept in mind, we do note that this final response pattern is consistent with Pennycook et al.'s (2015) original base-rate manipulation findings (i.e., less conflict detection with moderate base-rates).

General discussion

In this study, we tested three predictions of a hybrid dual process model in which people's intuitive reasoning performance is assumed to be determined by the absolute and relative strength of different intuitions (e.g., Bago & De Neys, 2017a; Pennycook et al., 2015). By manipulating the extremity of the base-rates in our reasoning problems we manipulated the strength of the logical intuition that is hypothesised to cue selection of the base-rate response. Consistent with the hybrid model predictions we observed that experimentally reducing the strength of the logical intuition decreased the number of initial (i.e., intuitive) base-rate responses when solving problems in which heuristic and logical intuitions conflicted. Second, reasoners who selected the stereotypical intuitive response were less likely to register the intrinsic conflict (as reflected in decreased confidence) in this case, whereas reasoners who selected the intuitive base-rate response experienced more conflict. These findings are hard to account for in the traditional serial default-interventionist (or parallel) model but provide support for the postulations of the hybrid dual process model.

The present study highlights how the hybrid dual process model generates new predictions that allow us to validate the model. Although we believe that the results illustrate the potential of the hybrid model view we also want to stress that the model is a "work in progress" and that there remain important challenges ahead. One issue concerns the specification of the role of System 2 processing in the framework. The general hybrid model that we put forward here focuses on the initial response stage in which System 2 is not activated. Although it postulates that detection of conflict will serve as a cue for the recruitment of System 2 (De Neys, 2012), it currently makes no further predictions about the nature of this System 2 processing. Interestingly, Pennycook et al. (2015) – one of the author teams that favored a hybrid model view – have attempted to provide such a further characterisation of the System 2 processing stage. In their three-stage model, the third processing stage specifies two different types of System 2 engagement that can follow intuitive conflict detection: Cognitive Decoupling and Rationalisation. Pennycook et al. explained the effects of their base-rate extremity manipulation primarily on the basis of these System 2 processes (e.g., less detection will lead to less rationalisation and hence, a less pronounced response latency increase). The present study highlights that similar effects can be observed on the basis of mere System 1 processing (note that this theoretical possibility was also recognised by Pennycook et al., 2015). During the critical initial response stage, System 2 processing was experimentally "knocked" out in the current study. Hence, in and by itself we do not need System 2 modulation to account for the effects of the base-rate extremity manipulation. To avoid confusion, it is important to stress that our findings do not argue against Pennycook et al.'s findings or model. The hybrid model that we put forward here focuses on the System 1 interaction between conflicting intuitions. It makes no further claims about the System 2 deliberation that might follow this conflict. In other words, it is possible that System 2 processes will result in similar effects. The model does not speak to this issue. The point is simply that the effects of a base-rate extremity manipulation can be observed in the absence of System 2 processing. System 2 processing might modulate these effects but it is not required to account for them. This implies that the precise role and possible unique contribution of System 2 remains to be specified in future work.

A second issue is that even the nature of the intuitive System 1 processing in the hybrid model (or any other dual process model for that matter) is in need of a more detailed specification. To illustrate, consider the recent findings of Bago and De Neys (2017b). In this study, we attempted to manipulate the strength of a heuristic intuition by changing the presentation order of the base-rates and descriptive information. Building on work of Pennycook et al. (2015) we hypothesised that whatever information is presented last would be more salient and increase in intuitive strength. Pennycook et al. found some evidence for this assumption with a classic single-response paradigm (e.g., presenting the description after – vs before – the base-rates decreased the number of base-rate responses). However, when we used the order manipulation with a two-response paradigm we observed the exact opposite effects at the initial response stage (Bago & De Neys, 2017b). This led Bago and De Neys to hypothesise that the last cued intuitive response had not reached its peak level at the enforced initial answer stage. As Bago and De Neys argued, although System 1 processing is assumed to be "fast" it is perhaps naïve to assume that intuitions are generated instantly at full strength. We need to factor in that they need some time to reach their peak strength (and will subsequently also decay in strength with the passing of time). The findings of Bago and De Neys (2017b) help us to start sketching a more fine-grained specification of the intuitive response generation mechanism. But the point we want to highlight here is that none of these features (i.e., rise and decay time) were a priori predicted by the hybrid model. Hence, arriving at a fully specified model of the postulated logical and heuristic intuition generation in the hybrid model will undoubtedly need further explorative work in the coming years.

A related question is what exactly constitutes the "strength" of an intuition. The hybrid model we proposed here uses "strength" as a general functional label to refer to the hypothesised activation level of an intuitive response. But "strength" and "activation level" can be operationalised in various ways (e.g., processing "fluency" or "speed"). At present the specific underlying processing specification and physical implementation remains to be characterised. Although we believe it is reasonable to rely on functional descriptions in theory development, we readily acknowledge that pinpointing the precise implementation remains an important challenge.

We noted that the "hybrid" model we presented here was inspired by the work of various scholars (e.g., Bago & De Neys, 2017a; Ball et al., 2017; Banks, 2017; Banks & Hope, 2014; Białek & De Neys, 2017; De Neys, 2012; Pennycook, 2017; Pennycook et al., 2015; Thompson & Newman, 2017; Trippas & Handley, 2017). Although we focused on the communalities, one might wonder about the precise relation between the various models we subsumed under the "hybrid" view. To avoid confusion, it might be worthwhile to explicitly point to some key developments. One early starting point for the hybrid framework was the "logical intuition" (De Neys, 2012) model. The model introduced the idea that System 1 cues both a heuristic and logical intuition which allowed to account for conflict detection findings and the evidence against the "bias blind" spot. Pennycook et al.'s (2015) "three-stage" model presented a more advanced hybrid model view that allowed to explicitly account for possible conflict detection failures while it also specified different types of System 2 engagement. Critically, Pennycook et al. were the first to explicitly postulate that differences in activation strength (i.e., generation speed in Pennycook et al.'s conceptualisation) might underlie detection failures: For some reasoners the logical intuition can be so weak that they will not register conflict with the stronger heuristic intuition. Bago and De Neys (2017) further built on this strength variability idea to account for the observation that some reasoners generated the appropriate logico-mathematical response intuitively in their two-response studies. Hence, Bago and De Neys, specified that logical intuitions can also dominate heuristic ones (a theoretical possibility that was also recognised by Pennycook et al., 2015). In sum, although these different models are constructed around a shared hybrid core, it should be clear that the latest version (i.e., Bago & De Neys, 2017a, 2017b) specifies features that were not specified in the initial De Neys (2012) version. For example, although De Neys' (2012) proposal entailed that different intuitions can have different strengths, it did not predict explicitly that there would be cases in which the logical intuition would dominate (or cases in which it would be absent, e.g., Pennycook et al., 2015, for that matter). But these observations are readily accounted for in the Bago and De Neys (2017) model or the Pennycook et al. (2015) model. This again illustrates the point that the hybrid view is a work in progress. It has been further developed and specified over the last couple of years and will need further specification and development in coming years. Therefore, it is critical to derive new predictions from the model and test these. It is here that the key contribution of the present paper lies.

Finally, one might also wonder how one needs to conceive the relation between the hybrid and traditional DI dual process model. Clearly, the hybrid model combines key features of the traditional serial DI model and parallel model (hence, it's "hybrid" name): Just like the serial model it assumes that System 2 processing is optional and starts later than System 1 processing. And just like the parallel model, it assumes that there is parallel logical and heuristic processing. However, unlike the parallel model it is claimed that this logical processing results from System 1 activation. Nevertheless, the hybrid model maintains the core DI assumption that people rely by default on System 1 processing and only switch to System 2 processing in a later stage of the reasoning process. That is, the hybrid model still maintains the DI feature that some initial System 1 processing always precedes System 2 processing (Bago & De Neys, 2017a; De Neys, 2014).

Note that Evans (2017) recently indicated that the traditional DI model also allows for the incorporation of logical intuitions. Evans' point is that popular reasoning and decision making tasks that have been used to test the hybrid view entail fairly simple logic rules and principles[6]. Therefore, these rules or principles might have become automatised through, for example, schooling and/or repeated exposure in daily life. Note that such an automatisation is precisely what De Neys (2012, 2014) has sketched as potential origin of the logical intuitions (see also Stanovich, 2018, for an interesting integrative perspective on automatisation and the degree of "mindware" instantiation).

In sum, the critical contribution of the present paper is that it demonstrates how the hybrid dual process model view – just like the traditional DI model in the past – allows us to derive new predictions that we can verify empirically. We believe that the current set of findings would be hard to account for in the traditional DI (or parallel) model and thereby lend credence to the hybrid model view. However, there is little reason to be triumphal. Even the hybrid model is still in its infancy. A key challenge will be to provide a more fine-grained specification of the nature of the different System 1 intuitions and the role of System 2 deliberation. At the same time, the present findings further underline the potential of a hybrid model view. In our opinion, it presents the most promising way forward for the dual process field.

Acknowledgements

Bence Bago was supported by a doctoral fellowship from Ecole des Neurosciences de Paris Ile-de-France. Open data statement: raw data can be retrieved from https://osf.io/k6tv2/?view_only=73be9bf8aae84c668b482815945e4371.

[6]For clarity, we note that the tasks that have been used to test the hybrid model (e.g., bat-and-ball problem, conjunction fallacy, ratio bias, base-rate neglect, belief bias syllogisms, etc) are the same tasks that have frequently been used to validate the DI model. We agree with the claim that these tasks are simple (De Neys, 2012). Indeed, problem complexity has been conceived to be a key boundary condition for logical intuitions (De Neys, 2012, 2014; Trippas, Thompson, & Handley, 2017). But we disagree with a possible implied suggestion that the tasks would therefore not be valid to test the DI predictions. The very same tasks have been used to argue in favor of the bias blind spot and corrective DI assumptions in the past (e.g., Kahneman, 2011; Stanovich & West, 2000).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was also supported by a research grant (DIAGNOR, ANR-16-CE28-0010-01) from the Agence National de la Recherche.

ORCID

Bence Bago http://orcid.org/0000-0001-6905-1832

References

Aczel, B., Szollosi, A., & Bago, B. (2016). Lax monitoring versus logical intuition: The determinants of confidence in conjunction fallacy. *Thinking & Reasoning*, *22*, 99–117. doi:10.1080/13546783.2015.1062801

Alós-Ferrer, C., & Strack, F. (2014). From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, *41*, 1–11.

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

Bago, B., & De Neys, W. (2017b). The rise and fall of conflicting intuitions during reasoning. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 87–92). Austin, TX: Cognitive Science Society. Retrieved from https://mindmodeling.org/cogsci2017/papers/0028/index.html

Bago, B., & De Neys, W. (2018). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking and Reasoning*. doi:10.1080/13546783.2018.1507949

Ball, L., Thompson, V., & Stupple, E. (2017). Conflict and dual process theory: the case of belief bias. In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.

Banks, A. (2017). Comparing dual process theories: evidence from event-related potentials. In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.

Banks, A. P., & Hope, C. (2014). Heuristic and analytic processes in reasoning: An event-related potential study of belief bias. *Psychophysiology*, *51*, 290–297.

Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). Reasoned connections: A dual-process perspective on creative thought. *Thinking & Reasoning*, *21*, 61–75.

Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, *12*, 148–167.

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition*, *38*, 186–196.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 356–366.

Cassotti, M., Agogué, M., Camarda, A., Houdé, O., & Borst, G. (2016). Inhibitory control as a core process of creative problem solving and idea generation

from childhood to adulthood. *New Directions for Child and Adolescent Development*, *2016*, 61–72.
De Neys, W. (2012). Bias and Conflict: A Case for Logical Intuitions. *Perspectives on Psychological Science*, *7*, 28–38.
De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*, 169–187.
De Neys, W. (2017). Bias, conflict, and fast logic: Towards a hybrid dual process future? In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.
De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PloS One*, *6*, e15954
De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299.
De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, *20*, 269–273.
De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load. *Experimental Psychology (Formerly Zeitschrift Für Experimentelle Psychologie)*, *54*, 128–133.
De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*, 483–489.
De Neys, W., & Verschueren, N. (2006). Working memory capacity and a notorious brain teaser: The case of the Monty Hall Dilemma. *Experimental Psychology*, *53*, 123–131.
Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *The American Psychologist*, *49*, 709–724.
Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.
Evans, J. (2010). *Thinking Twice: Two Minds in One Brain*. Oxford: Oxford University Press.
Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science*, *8*, 223–241.
Evans, J. (2017). Dual Process Theories: Perspectives and problems. In W. De Neys (Ed.), *Dual Process Theory 2.0* (pp. 137–155). Oxon, UK: Routledge.
Franssens, S., & De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking & Reasoning*, *15*, 105–128.
Gangemi, A., Bourgeois-Gironde, S., & Mancini, F. (2015). Feelings of error in reasoning—in search of a phenomenon. *Thinking & Reasoning*, *21*, 383–396. doi:10.1080/13546783.2014.980755
Gilhooly, K. J. (2016). Incubation and intuition in creative problem solving. *Frontiers in Psychology*, *7*, 1076.
Greene, J. (2013). *Moral tribes: Emotion, reason and the gap between us and them*. New York, NY: Penguin Press.
Gürçay, B., & Baron, J. (2017). Challenges for the sequential two-system model of moral judgement. *Thinking & Reasoning*, *23*, 49–80.
Johnson, E. D., Tubau, E., & De Neys, W. (2016). The Doubting System 1: Evidence for automatic substitution sensitivity. *Acta Psychologica*, *164*, 56–64.
Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
Mata, A., Ferreira, M. B., Voss, A., & Kollei, T. (2017). Seeing the conflict: An attentional account of reasoning errors. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-017-1234-7

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences*, *22*, 280–293.

Mevel, K., Poirel, N., Rossi, S., Cassotti, M., Simon, G., Houdé, O., & Neys, W. D. (2015). Bias detection: Response confidence evidence for conflict sensitivity in the ratio bias task. *Journal of Cognitive Psychology*, *27*, 227–237. doi:10.1080/20445911.2014.986487

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.

Newman, I., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief -bias and base-rate neglect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1154–1170.

Pennycook, G. (2017). A perspective on the theoretical foundation of dual process models. In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014). Cognitive style and religiosity: The role of conflict detection. *Memory & Cognition*, *42*, 1–10.

Pennycook, G., De Neys, W., Evans, J. S. B. T., Stanovich, K. E., & Thompson, V. (2018). The mythical dual-process typology. *Trends in Cognitive Sciences*, *22*, 667–668.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning? *Cognition*, *124*, 101–106.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72.

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 544–554.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489*, 427–430.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science*, *13*, 60–66.

Simon, G., Lubin, A., Houdé, O., & De Neys, W. (2015). Anterior cingulate cortex and intuitive bias detection during number conservation. *Cognitive Neuroscience*, *6*, 158–168. doi:10.1080/17588928.2015.1036847

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.

Stanovich, K. (2011). *Rationality and the reflective mind*. Oxford: Oxford University Press.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*. Advanced Online Publication

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, *23*, 645–665.

Stupple, E. J., Ball, L. J., & Ellis, D. (2013). Matching bias in syllogistic reasoning: Evidence for a dual-process account from response times and confidence ratings. *Thinking & Reasoning*, *19*, 54–77.

Stupple, E. J., Ball, L. J., Evans, J. S. B., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology*, *23*, 931–941.

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, *20*, 215–244.

Thompson, V. A., Turner, J. A. P., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*, 107–140.

Thompson, V., & Newman, I. (2017). Logical intuitions and other conundra for dual process theories. In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, *150*, 109–118.

Trippas, D., & Handley, S. (2017). The parallel processing model of belief bias: review and extensions. In W. De Neys (Ed.), *Dual Process Theory 2.0*. Oxon, UK: Routledge.

Trippas, D., Thompson, V. A., & Handley, S. J. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition*, *45*, 539–552.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science (New York, N.Y.)*, *185*, 1124–1131.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1310–1321.