Thinking & Reasoning

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ptar20

Uncontrolled logic: *intuitive sensitivity to logical structure in random responding*

Stephanie Howarth, Simon Handley & Vince Polito

To cite this article: Stephanie Howarth, Simon Handley & Vince Polito (2022) Uncontrolled logic: *intuitive sensitivity to logical structure in random responding*, Thinking & Reasoning, 28:1, 61-96, DOI: 10.1080/13546783.2021.1934119

To link to this article: https://doi.org/10.1080/13546783.2021.1934119

Published online: 22 Jun 2021.

THINKING & REASONING
2022, VOL. 28, NO. 1, 61–96
https://doi.org/10.1080/13546783.2021.1934119

Uncontrolled logic: *intuitive sensitivity to logical structure in random responding*

Stephanie Howarth, Simon Handley and Vince Polito

Department of Cognitive Science, Macquarie University, Sydney, Australia

ABSTRACT
It is well established that beliefs provide powerful cues that influence reasoning. Over the last decade research has revealed that judgments based upon logical structure may also pre-empt deliberative reasoning. Evidence for 'intuitive logic' has been claimed using a range of measures (i.e. confidence ratings or latency of response on conflict problems). However, it is unclear how well such measures genuinely reflect logical intuition. In this paper we introduce a new method designed to test for evidence of intuitive logic. In two experiments participants were asked to make random judgments about the logical validity of a series of simple and complex syllogistic arguments. For simple arguments there was an effect of logical validity on random responding, which was absent for complex arguments. These findings provide a novel demonstration that people are intuitively sensitive to logical structure.

ARTICLE HISTORY Received 18 October 2020; Accepted 20 May 2021

KEYWORDS Reasoning; intuitive logic; dual process theory; syllogistic reasoning; random responding

Introduction

One of the most important and influential claims of recent decades concerning the cognitive architecture underlying human reasoning is that our judgments reflect the operation of two distinct 'systems' or 'types' of processing. This dichotomy is captured by the classic Dual Process (DP) theory, a metatheoretical framework that pitches heuristic or intuitive (Type 1) thinking processes, against deliberative or analytical (Type 2) processes (Stanovich & Toplak, 2012). Evans and Stanovich (2013), have argued that

CONTACT Stephanie Howarth stephanie.howarth@mq.edu.au; Simon Handley simon.handley@mq.edu.au Department of Cognitive Science, Macquarie University, Sydney NSW2109, Australia.

Supplemental data for this article is available online at https://doi.org/10.1080/13546783.2021.1934119.
All supplementary materials are accessible via the Open Science Framework website at: https://osf.io/qzsmk/?view_only=12dc8d3028b34f309f9d9fe3b581b38e

the defining feature of Type 1 processes is that they are autonomous and independent of capacity constraints, whereas Type 2 processes are controlled and dependent upon working memory resources. Dual process frameworks have been applied across a diverse range of fields including memory (Smith & DeCoster, 2000), moral reasoning (Greene, 2013), delusional beliefs (Coltheart et al., 2011), decision biases (Kahneman, 2011), cognitive development (Barrouillet, 2011), cognitive neuroscience (Goel, 2003) and learning (Reber, 1996).

One of the most influential DP models within the reasoning and decision-making domain, is the default-interventionist (DI) serial framework (Evans, 2003; Evans & Stanovich, 2013). The proponents of this model claim that Type 1 heuristic responses to reasoning and judgment tasks are typically based on learned associations or beliefs, and are initiated quickly and generated by default. Effortful, Type 2 processes must be engaged in order to inhibit and override the often-incorrect heuristic output. In other words, default responses are generated automatically when people reason or make a decision, and these are typically accepted because intervention and override require cognitive effort. Furthermore, our inclination towards miserly processing (Toplak et al., 2014), means that, even with the appropriate 'mindware' to support an accurate response (Stanovich, 2018), most of us fail to exert the effort required to intervene on a readily available heuristic response.

The Belief Bias effect (Evans et al., 1983) is an example where default beliefs impair a reasoner's ability to make accurate deductive judgments. It is one (amongst many) cognitive biases that have been claimed to provide support for the default interventionist DP account of reasoning. Over the decades, there has been much research that supports a default interventionist interpretation of the Belief Bias effect (Evans, 2003, 2006, 2008). For example; studies designed to minimise Type 2 processing by increasing cognitive load (De Neys, 2006), imposing time restrictions (Evans & Curtis-Holmes, 2005) or using instructional manipulations (Evans et al., 2010), have been shown to increase the Belief Bias effect. Conversely, using reasoning problems with emotionally charged negative content, that violates social norms, reduces Belief Bias (Goel & Vartanian, 2011), as does the provision of elaborated instructions of logical principles (Evans et al., 1994).

Within the last decade, however, research has shown that, in some instances, judgments based upon beliefs require effort and take longer to process than judgments based upon logical validity. Handley et al. (2011) demonstrated this using an instructional manipulation paradigm, which required participants to respond to a set of reasoning questions, under typical logic instructions (does the conclusion logically follow?) or under belief instructions (is the conclusion believable or unbelievable?). Contrary to the predictions of the Default Interventionist account they showed that the presence of a conflict between the validity and believability of a conclusion, had a greater impact on the accuracy of belief judgments than logic judgments, resulting in more errors under belief instruction particularly when inhibitory demands are increased (Howarth et al., 2016; see also Howarth et al., 2019). These findings are difficult to reconcile with the classic DI account, which assumes that belief activation and responding is based upon an effortless and quick, Type 1 process.

In conjunction with evidence of effortful belief judgments (see Handley & Trippas, 2015, for a review), a growing number of studies have claimed to have identified instances of effortless logic, commonly referred to as 'logical intuition'. The idea that people possess a form of intuitive logic, is not in itself new. Theories of natural deduction have maintained that certain inferences are triggered automatically in the process of understanding logical connectives (Braine & O'Brien, 1991; Rips, 1994), and the evidence that logical validity interferes with conflicting belief judgments, described above, is consistent with the idea that certain logical inferences are available rapidly.

This raises the question of whether the observation of a 'biased' response on a reasoning task arises because the biased judgment pre-empts logical processing, or whether the logical response is simultaneously available but simply not selected? In other words, do people know they are biased when they reason? The conflict detection paradigm was introduced to evaluate whether reasoners detect conflict, even if they generate a biased response, as this would offer evidence that people simultaneously process a problem's logical structure. A number of behavioural findings have provided strong evidence that the majority of reasoners are implicitly aware of conflict irrespective of the response that they generate. For example, people take longer to respond to conflict compared to no-conflict problems (De Neys, 2012; De Neys & Glumicic, 2008), they also inspect them for longer (Stupple & Ball, 2008), report lower confidence ratings (De Neys et al., 2011) and higher feelings of error when attempting to solve them (Gangemi et al., 2015). Studies conducted under time pressure or cognitive load (De Neys, 2017) have confirmed that conflict sensitivity occurs at an implicit level, which is further supported through physiological measures (De Neys et al., 2010) and increases in activation in areas of the brain (the Anterior Cingulate Cortex) associated with the error and conflict detection (De Neys et al., 2008).

The research offers persuasive evidence that sensitivity to conflict indicates some processing of logical structure at an intuitive level. Furthermore, for conflict to occur, the processing of logical structure must occur simultaneously and in parallel with heuristic/belief-based processing. Banks and

Hope (2014) lend support to this view with EEG data showing a heightened P3 positivity for conflict problems irrespective of whether a belief or logic response was generated. This implies that logic and beliefs both impact reasoning early and concurrently (also see Bago et al., 2018).

Further support for the claim of logical intuition comes from recent research using a two-response paradigm (Thompson et al., 2011). This method requires participants to provide a rapid response, typically within a restricted amount of time. They are then asked to respond again but with time to reflect on their final answer (Bago & De Neys, 2017; Newman et al., 2017). This method has shown that those who gave the correct response after deliberation, also gave the correct response at the intuitive stage, and did so with high confidence (Bago & De Neys, 2017). This suggests that the logical response is available early and is rarely modified following further deliberation.

This corroborates the work using the instruction manipulation paradigm described earlier, where participants are instructed to make a judgement based on logical validity (Handley et al., 2011), or base-rate information (Pennycook et al., 2014); or on the basis of belief (conditionals/syllogisms) or the description (accompanying the base-rate task). Findings from these studies demonstrate that logical (deductive inferences) and statistical processing (base rate task) are accomplished rapidly and interfere with belief-based judgments. Both paradigms, therefore, support the notion of 'intuitive logic' which may be the default response, processed outside of conscious awareness or deliberation.

Up to this point, much of the evidence substantiating instances of intuitive logic tends to rely on procedures designed to eliminate any Type 2 processing, such as the use of speeded tasks or increased cognitive loads. One issue with these techniques is that it is difficult to determine definitively that they eliminate the capacity for reasoners to engage in some form of deliberative thinking. Recent research has shown that participants of higher cognitive capacity tend to show 'intuitive logic' effects (Raoelison et al., 2020; Thompson et al., 2018), but as higher cognitive capacity translates into greater processing speed (Fry & Hale, 1996), it could equally be the case that increased capacity supports the engagement of deliberative processing even under speeded conditions. In addition, recent evidence suggests that some reasoners are able to develop the ability to make rapid judgments that depend upon complex mathematical calculations. For example, Raoelison and De Neys (2019) revealed that with repeated exposure to variations of the Cognitive Reflection Task, using the two-response paradigm, a small number of individuals were able to 'learn' how to generate the correct response at the initial stage on subsequent trials. At the very least, this suggests that some participants are able to adapt their processing strategy to support complex structural processing accomplished under demanding time constraints.

Other researchers, have explored alternative more indirect methods of investigating intuitive contributions to reasoning. The 'logic–liking' technique was introduced by Morsanyi and Handley (2012), based on work by Whittlesea (1993) and later Topolinski and Strack (2008, 2009) which used word triads and liking ratings to explore intuitive coherence judgments. Applied to classic syllogistic reasoning problems, they asked people to simply rate how much they liked the concluding statements of logical arguments, as a way of measuring the affective response elicited by valid and invalid arguments. The logic–liking paradigm is built around the idea that valid arguments are processed more fluently and this conceptual fluency gives rise to positive feelings which are construed as more likable. The novelty with this method is that the task makes no reference to logical reasoning, yet people consistently showed sensitivity to logical structure with higher liking ratings for conclusions that follow validly from the preceding statements compared to those that do not. This effect has been replicated both with conditional inferences (Trippas et al., 2016) and syllogistic arguments. Recently, however, research has called in to question the extent to which liking judgments reflect intuitive sensitivity to logical structure showing that the validity effect in liking judgments is related to cognitive ability and can be reduced by working memory load (Hayes et al., 2020). This suggests that the validity effect on liking judgments may arise as a result of explicit deliberative reasoning, which is more effective amongst participants of higher cognitive capacity or those with more available working memory capacity. It may well be the case that some participants engage in explicit reasoning in seeking relevant evidence to underpin their liking judgments.

In the pursuit of a purer measure of intuitive logic, Trippas et al. (2016) used a technique that made no explicit reference to logicality but instead asked people to rate the 'brightness' (contrast) of the conclusion to various types of logical argument. They showed that people rated valid statements as brighter than sentences that did not follow logically from preceding premises, which again was explained within the framework of fluency misattribution, whereby positive affect is elicited by processing fluency (see Nakamura & Kawaguchi, 2016; Winkielman & Cacioppo, 2001). In this respect, valid arguments create a greater processing fluency, which produces positive feelings that are ambiguous and hard to interpret due to lack of insight, whereas invalid arguments are associated with negative feelings arising from dysfluency. These feelings are then misattributed to higher (or lower) brightness ratings.

These findings demonstrate that the validity of an argument can influence judgments that require the evaluation of features that are entirely independent of an argument's underlying structure. They are consistent with the idea of 'uncontrolled intelligence', a concept initially described by Wegner et al. (2003) who examined instances of apparently intelligent action without conscious intention or awareness. Wegner et al. were motivated by research into a technique called Facilitated Communication (FC), which was developed to support people with communication disabilities to communicate through assisted typing. The method involves facilitators bracing their client's hands at a keyboard whilst they type. Although facilitators claim not to contribute to the messages produced, research has revealed that many of the responses originate from the facilitators rather than their clients. Facilitators are often convinced that communication was initiated from the patient, leading Wegner to argue that they are exhibiting intelligent action without any conscious awareness or sense of agency. In a series of studies, Wegner et al. (2003) explored the idea that prior knowledge could influence action, against intention. To do this they asked people to generate random yes and no responses to a set of easy and hard trivia questions, in order to investigate whether people's knowledge could unintentionally inform their ability to give random responses. They found that participants responses were reliably above chance levels on easy trivia questions, but not on more difficult ones, confirming that they were influenced by their knowledge of the accuracy of the answers, but only when the solution was readily available. Furthermore, participants were unaware that their responses were being influenced by relevant knowledge of response veracity, demonstrating a lack of insight into the basis of their responding. Manipulations that are typically considered to increase cognitive control (time limits and financial incentives) had no impact on performance in the clever hands task, leading Wegner et al. to hypothesise that knowledge of the correct answers influenced responses in an automatic and uncontrolled fashion. These findings have since been replicated by Polito et al. (2018), and extended to highly hypnotisable participants under a specific suggestion to respond randomly.

In this article we are following up on Wegner's work, using his 'random response' instructional method, to evaluate whether people will be influenced by the logical validity of an argument when asked to respond randomly on a judgment task. Our intention was to test the 'intuitive logic' hypothesis using a novel method which makes no explicit reference to the logical features of the task, but rather instructs participants to respond randomly, regardless of any structural or presentational features of the problem. This represents a strong test of the intuitive logic hypothesis. Under Wegner's analysis, if responses are influenced by the validity of the stimuli, this must arise from the automatic impact of 'logical knowledge' on action, without intention.

Across two experiments we instructed participants to complete a set of simple (Experiment 1) and complex (Experiment 2) syllogistic reasoning problems and answer according to logic or to respond randomly. In Experiment 1, we predicted that for simple syllogisms, despite the instruction to respond randomly, the logic of the argument would be processed automatically and rapidly, leading to higher rates of endorsement for valid than invalid arguments. We also manipulated the believability of the conclusion in both experiments. Based on previous research that has shown that the logical validity of an argument has a pre-emptive impact on belief judgments, we predicted that the logical features of the argument's conclusion, rather than its believability, would have the most significant impact on judgments. In Experiment 2, we predicted that increasing the complexity of the arguments would mitigate intuitive logic effects, and did not expect differences in endorsement rates between valid and invalid arguments. We also collected post task estimates of accuracy and randomness, to evaluate the extent to which participants had insight into the basis for their responding. Finally, we administered the Cognitive Reflections Task to determine whether there was a relationship between cognitive style and responding under the two instructional conditions.

Experiment 1

In Experiment 1, we presented participants with simple 'single-model syllogisms' (see Johnson-Laird & Byrne, 1991; Trippas et al., 2013, 2017) in two blocks, one under logic instruction, and one using the novel random instruction method (Wegner et al., 2003), presented in counterbalanced order. Our aim was to determine whether the logical structure of an argument would influence judgments under the instruction to respond randomly or whether responding randomly was independent of logical validity.

Method

Participants

Fifty-five undergraduates were recruited through the Psychology SONA system at Macquarie University and received course credit for participating in the Experiment. One individual was eliminated for not completing the experiment. This left a total of fifty-four females and nine males (mean = 20 yrs.).

Design, materials & procedure

A 2 (Validity: Valid/Invalid) x 2 (Believability: Believable/Unbelievable) x 2 (Instruction: Logic/Random) fully repeated measures design was used,

Table 1. Examples of simple syllogism for Experiment 1.

Valid		Invalid	
Believable	Unbelievable	Believable	Unbelievable
All meetals are cars Some Mazdas are meetals Some cars are Mazdas	All meetals are cars Some Mazdas are meetals Some Mazdas are cars	Some roses are crinds All crinds are birds No roses are birds	Some roses are crinds All crinds are flowers No flowers are roses

where each participant was presented a total of 64 one-model type syllogisms on a computer screen using the E-prime software.

The one model structures were taken from Trippas et al. (2013, 2017), using all quantifiers (All, No, Some) except 'Some … not' and sixteen syllogistic structures (see supplementary materials A, for the syllogism structures). The content of the syllogisms were a mix of real and nonsense words created using the pseudo word generator Wuggy (Keuleers & Brysbaert, 2010). Half the syllogisms were conflict problems and half were no-conflict problems. The conflict problems consisted of 8 Valid-Unbelievable and 8 Invalid-believable items, whilst the no-conflict problems consisted of 8 Valid-Believable and 8 Invalid-Unbelievable items.

The content of the syllogisms referred to categories and category members. There was a total of 32 categories and four possible members for each category (i.e. **Category**: criminal; **Member**: murderers, thieves, kidnappers, terrorists). The stimuli were presented in 2 blocks under different instructions and order was counterbalanced between participants. To ensure there would be no repetition of content across blocks, the stimuli was pseudo randomised; where the 16 categories presented in block 1 were distinct from the 16 categories presented in block 2. In each block, categories were presented twice with different members to make up 32 items in a block. Categories were counterbalanced across instruction type, and category/member combinations were randomly allocated to one of the syllogistic structures. This ensured that each participant was presented with a unique list of 64 items in total (see Table 1 for an example of the items used).

Instructions. One block was presented under logic instruction and one block under random instruction. Under logic instruction participants were presented with the following instructions:

> In this part of this experiment, we are interested in your ability to make judgments on the basis of LOGIC.
>
> You should assume all the information presented is true
> (even if it's not, or if it doesn't appear to make much sense).
>
> The first and second premise will be displayed first for a few seconds,
> then shortly after, a concluding sentence BELOW the line will appear,
> which you will be asked about.

> If you judge that the conclusion necessarily follows from the premises,
> you should answer 'Valid' by pressing the 's'-key,
> otherwise you should answer 'Invalid' by pressing the 'k'-key.
>
> For example:
>
> All cars are blurbs
> All blurbs are cheap
>
> All cars are cheap
>
> On the basis of LOGIC, the correct response is 'Valid',
> because the sentence 'All cars are cheap' necessarily follows from
> the premises above the line (if you assume they are true).

In the random block of trials, participants were presented with the following instructions:

> In this part of the experiment, we are interested in your ability to be RANDOM.
> You will be presented with a set of trials in the following format:
>
> All cars are blurbs
> All blurbs are cheap
>
> All cars are cheap
>
> Two premises presented for a short time followed by a concluding sentence BELOW the line.
> Response options are 'Valid' by pressing the 's'-key
> or 'Invalid' by pressing the 'k'-key.
>
> Your aim is to answer Valid or Invalid as RANDOMLY as possible,
> regardless of the validity of the conclusion.
>
> Try NOT to generate a predictable pattern of 'Valid, Valid, Valid' responses or thus like,
> but do try and generate random sequences.
> For each trial, try and make the most FREE and RANDOM choice you possibly can.

Premise 1 was presented for 1000 ms alone on the screen, the second premise was then added for a further 1000 ms, followed by the conclusion and the two response options. The full problem remained on the screen until the participant gave a response.

In order to ensure that the participants were reading and comprehending the statements presented in the random block, we informed the participant of the following;

> Remember to read and pay attention to each trial and its content
> because you will be asked to complete a
> MEMORY TEST afterwards.
>
> You will be asked to identify whether certain words had been
> presented throughout this block of trials
> so stay alert.

Just before the memory block they were reminded of the following;

This part of the experiment is the
MEMORY TEST.

You will be presented with words one at a time and you have to indicate whether they were present in the previous block of trials by pressing the 's'-key for YES
the 'k'-key for NO.

The memory test consisted of 32 words, half of which were from the content of the random block of trials and half were not.

Post block estimates

After each block, participants were asked to make a judgment about their performance on the task as a way of determining whether people were subjectively aware of the influence of logic or beliefs on their responding. Participants were required to estimate the percentage of questions they thought they answered correctly on the logic and random blocks of trials as a measure of 'estimated correctness' and evaluate the randomness of their performance on the random block, as a measure of 'estimated randomness'. Estimations were measured on a scale of 1 to 100 with 100 being 'completely random' or 'completely correct'.

Individual differences measure

At the end of the experiment, participants completed the 7-point Cognitive Reflection Test (CRT) as a measure of cognitive style (Toplak et al., 2014).

Results

Our main analysis included four separate repeated measures ANOVAs on endorsement rates (number of 'Yes' responses) and latencies for both logic and random instructions, using a 2 (Validity: Valid/Invalid) × 2 (Believability: Believable/Unbelievable) design. We did not include task order in the analysis because this was counterbalanced, however, given that completing a block of logic problems prior to a block of problems under random instruction might facilitate sensitivity to logical structure, where appropriate we do report the magnitude of the interaction between order and the effect of validity. The full analysis, including order as a variable is included in supplementary materials B. Response latencies (as measured from the presentation of conclusion to response) were Log transformed to approximate a more normal distribution better aligned with the assumptions of the ANOVA. The untransformed latencies are presented in Table 2 for ease of interpretation.

Table 2. Endorsement rates and raw latencies, across problem type under Logic and Random Instruction.

	Valid		Invalid	
Instruction	Believable	Unbelievable	Believable	Unbelievable
Logic – Endorsements (%)	**90(12.8)**	**82(22.8)**	**17(26.6)**	**14(18.9)**
Logic – Latencies (ms)	**9,584(5058)**	**10,188(4582)**	**10,097(4859)**	**9,945(5209)**
Random – Endorsements (%)	**59(19.8)**	**53(20.6)**	**47(21.0)**	**46(19.3**
Random – Latencies (ms)	**4,038(3994)**	**4,492(4591)**	**4,311(4373)**	**4,388(4231)**

SD in brackets.

Before conducting the main analysis, we examined the memory test scores to ensure that participants were sufficiently engaged during the random block trials. Accuracy was high with participants making correct recognition judgments on average 83% of the time; One participant was eliminated from the analysis who scored more than 2SDs below the sample mean (below 62% accuracy), giving a total *n* of 53 for the analyses that follow.

Logic instruction

A 2 (Validity: Valid/Invalid) by 2 (Believability: Believable/Unbelievable) within participants analysis of variance on endorsement rates showed a large main effect of *validity*; $F(1, 52) = 257.944$, MSE $= 1032.1$, $p < .001$, $\eta_p^2 = .832$, where valid conclusions were endorsed more than invalid conclusions (86% vs. 15%) and a small main effect of *believability*; $F(1, 52) = 4.081$, MSE $= 365.7$, $p = .049$, $\eta_p^2 = .073$, where believable conclusions were endorsed more than unbelievable conclusions (53% vs. 48%). There was no interaction between *validity* and *believability*; $F(1, 52) = 3.733$, MSE $= 104.4$, $p = .059$, $\eta_p^2 = .067$.

A similar analysis of variance carried out on the log transformed response latencies showed no main effect of *validity*; ($F < 1$), no effect of *believability* ($F < 1$) and no significant interaction between *validity* and *believability;* $F(1, 52) = 3.690$, MSE $= .010$, $p = .060$, $\eta_p^2 = .066$.

Random instruction

The 2 (Validity: Valid/Invalid) by 2 (Believability: Believable/Unbelievable) within participants analysis of variance on endorsement rates under random instructions showed a main effect of *validity*; $F(1, 52) = 5.650$, MSE $= 847.5$, $p = .021$, $\eta_p^2 = .098$, indicating that participants endorsed more valid than invalid conclusions (56% vs. 47%). There was no main effect of *believability*; $F(1, 52) = 2.730$, MSE $= 301.7$, $p = .105$, $\eta_p^2 = .050$, and the interaction between validity and believability was not significant ($F < 1$). The effect of validity shows that even under instructions to respond randomly logical structure continued to have a systematic influence on responding. Further analysis suggests that this effect extends across a majority of the sample, with 63% of participants showing accuracy rates above 50%. Interestingly, and in line with the ANOVA analysis, a much smaller proportion of participants (45%) had more than 50% of their responses aligned with the believability of the conclusion, despite this being a potentially more salient cue to responding.

To ensure that the observed effect of validity on random responding did not arise due to participants being cued to the logical features of the task after completing the logical block first, we reran the analysis including order as a between subjects' factor. There was no interaction between validity and order (F (1,51) = 2.27, MSE = 827, p = .14, η_p^2= .043), indicating that the effect of validity did not depend upon participants completing the logic block first. In fact, the size of the validity effect was larger in absolute terms when the random block was presented first (M_v= 59, M_{inv}= 43) rather than following the logic instruction block (M_v= 53, M_{inv}= 49).

The analysis of response latencies under random instructions showed no main effect of *validity* ($F < 1$) but there was a main effect of *believability;* F(1, 52) = 5.869, MSE = .014, p = .019, η_p^2= .101, showing that participants were quicker to respond to believable items (M = 4,175ms) compared to unbelievable items (M = 4,440ms). There was no interaction between these variables ($F < 1$). It is worth noting that latency of response under random instructions (M = 4,307ms) was significantly lower than under logic instructions (M = 9,954ms; t(52) = 7.88, p <.001), indicating that participants, as expected, were engaged for substantially less time when instructed to respond randomly compared to logically.

Task performance, self-insight and cognitive style

Our second set of analyses were designed to evaluate the extent to which participants showed insight into their responding, both in terms of its accuracy and, for the random instruction task, its randomness. An important question, which has generated significant recent debate, concerns the best way of measuring performance accuracy on reasoning tasks with a binary yes/no response. A typical approach might involve deriving a 'logic index' where the number of endorsements of invalid conclusions is subtracted from the number of endorsements of valid conclusions. An index of belief bias can similarly be calculated by subtracting unbelievable from believable endorsements. These indices are often used to evaluate how logical sensitivity and bias relate to other measures such as cognitive capacity, style or subjective judgments of performance.

Recently a number of researchers have questioned the assumptions underpinning this method as it assumes a linear relationship between Hits (saying valid when the argument is valid) and False alarms (saying valid when the argument is invalid), for different levels of response bias (reflecting where the criteria for 'yes' responding is set).

These authors have demonstrated that, in fact, the relationship between Hits and False Alarms for different response criteria is curvilinear and hence Signal Detection Theory (SDT) is a more appropriate framework for analysing belief bias data (Dube et al., 2010; Heit and Rotello, 2014). The theory focuses on signals (targets) and noise (non-targets) in a set of presented stimuli. For example, in this study there are *valid* (correct – signal) and *invalid* (incorrect – noise) response options. If an individual answers *valid* to a *valid* argument, this is considered a Hit (H). If they respond *valid* to an *invalid* argument, this is considered a False Alarm (F). From Hs and Fs you can calculate an individual's level of *sensitivity* (d'prime: $d' = z(H) - z(F)$) to the stimulus presented and *response bias* (Criterion $= c = -0.5 *[z(H) - z(F)]$). In the current example, *sensitivity* refers to an individual's ability to discriminate between valid and invalid arguments. Response bias helps us determine an individual's inclination to say yes (or valid) more or less than no (or invalid) and the difference in criterion setting for believable compared to unbelievable conclusions provides a measure of 'belief bias'.

Dube et al. have argued that reasoning data are best analysed using unequal variance SDT models and hence, in order to estimate sensitivity and bias accurately, it is necessary to simultaneously collect confidence data which allows for more accurate modelling and estimation of sensitivity and bias. However, a recent large meta-analysis, re-analysing 22 confidence ratings studies, showed that confidence rating data is largely unnecessary (Trippas et al., 2018) and that an equal variance SDT model is suitable for modelling the sensitivity and bias of syllogistic reasoning data, which means that these indices can be derived directly from Hits and False alarms, in the way described above.

It is this analytic approach that we employ in the following analyses, which examine the extent to which participants are able to demonstrate self-insight into their responding. One might expect, for example, if reasoners are unaware of the impact of validity on random responding, then sensitivity would be unrelated to judgments concerning 'correctness' of responding. A key question concerns how these patterns of relationship might vary across the instructional conditions.

We calculated the standard d-prime (d') as our index of senstivity to logical validity and a belief bias index reflecting the difference in criterion for unbelievable compared to believable conclusions (c_u-c_b). The analyses of the SDT indices confirmed the findings of the ANOVAs described above. Under random instructions sensitivity was significantly above zero ($M = .26$, $t(52) = 2.36$, $p = .02$). There was no reliable difference in bias between arguments with believable ($M = -.11$) and unbelievable conclusions ($M = .01$,

$t(52) = -1.73$, p =.09), although the effect would be significant under a one-tailed test, which could be justified based on the expectation of more conservative responding on unbelievable problems. Similarly, under logic instructions sensitivity was significantly above zero ($M = 2.4$, $t(52) = 16.03$, $p < .001$) and there was a significant one-tailed effect of beliefs ($M_{cb} = -.09$, $M_{cu} = .07$, $t(52) = 1.83$, $p = .036$, one-tailed).

Having confirmed the evidence of a logic effect under random instructions using our measure of sensitivity, the next set of analyses examined the extent to which reasoners demonstrated self-insight into their performance. Recall, that in line with Wegner et al., we asked participants to make post-task judgments about task accuracy under both instructional conditions and to judge how random they considered their responses to be in the random instruction condition. Participants made these judgments on a 100-point scale.

Interestingly, observed accuracy under random instructions ($M_o = 55$) was significantly higher than estimated accuracy ($M_e = 32$, $t(52) = 7.12$, $p <.001$), a pattern that also held under logic instructions ($M_o = 85$, $M_e = 55$, $t(52) = 10.35$, $p < .001$). However, there was no evidence that the difference between estimated and observed accuracy was any greater under random than logic instructions ($M_{dr} = 23$, $M_{dl} = 30$, $t = 1.69$, $p = .096$.). Curiously, estimates of accuracy under random responding where substantially below 50%, perhaps because participants judged that random responding would lead to very low rates of accuracy (instead of 50% accuracy). Importantly though, the accuracy estimates differed between instructional conditions, suggesting that participants were responding and judging their performance quite differently when responding randomly compared to when responding under logical instruction.

In order to determine whether participants had insight into the logical accuracy of their responding, we carried out a median split of estimated accuracy scores and compared sensitivity for participants who judged higher accuracy (High group) compared to those that judged lower accuracy (Low group). Under logic instructions the High group mean accuracy rating was 73 ($N = 28$) and the Low group ($N = 25$) was 35, but there was no accompanying difference between the groups in sensitivity ($M_H = 2.64$, $M_L = 2.15$, $t(51) = 1.62$, $p = .110$). Under random instructions, the High ($N = 32$, $M_e = 42$) and Low groups[1] ($N = 21$, $M_e = 16$) also showed no reliable difference in sensitivity ($M_H = .32$, $M_L = .18$, $t(51) = .60$, $p = .554$). We similarly considered the randomness estimates provided following the random instruction trials. Once

[1]For the measure of estimated correctness under random instruction 14 participants had estimates that were equal to the median. In these cases, they were allocated to the high group. The same method was applied to estimates of correctness under logic, where 10 participants had estimates that were equal to the median.

again there was no evidence that participants who judged their responses to be more random (N = 27, M_{er} = 82) compared to less (N = 26, M_{er} = 47) were any less influenced by the logical validity of the arguments as reflected in a comparison of sensitivity (M_H = .29, M_L =. 24, t(51) = .2, p = .842).

The analysis of the relationship between performance estimates and sensitivity suggest that participants have little insight into the impact of argument validity on responding under random instructions. However, based upon the analyses above, this pattern is not unique to random judgments and extends to the logic instruction condition, a finding aligned with observations in the literature that there is often little between subjective measures of performance, such as confidence or 'feelings of rightness', and logical accuracy (Thompson et al., 2011; Thompson & Morsanyi, 2012).

There is another way we can test the extent to which the validity effects under logic and random instructions arise through different mechanisms. In this study we also included a measure of cognitive style, the seven item Cognitive Reflection Task, which has often been shown to correlate with performance on reasoning tasks where there exists conflict between logical and belief based responses, presumably because the task indexes the degree to which participants are able to reason based upon the underlying logical structure, independent of its superficial content. Interestingly we observed a significant correlation between performance on the CRT and sensitivity under logic instructions (r(51) = .35, p = .01), a pattern that was reversed in the random condition (r(51) = −.21, p = .14). These correlation coefficients were significantly different as shown by a Fishers R to Z transformation (z = 2.88, p = .002). This finding provides some preliminary evidence that the two tasks are measuring different types of sensitivity to logical structure, a claim that is consistent with zero order correlation between the measures of sensitivity across the two instructional tasks (r = .096, p = .494). It is worth noting that the measures of sensitivity for both logic instructions (alpha = .81) and random instructions (alpha = .60) had acceptable levels of split half reliability.

Finally, we explored the extent to which logical sensitivity under the logic and random responding conditions could be predicted through a combination of the measures of estimated accuracy, CRT score and sensitivity on the other task. We carried out two multiple regression analyses using the enter method, the first of which was on the logic instruction condition with sensitivity as the dependent variable and CRT score, estimated accuracy and sensitivity under random instructions as the predictors. The overall model was significant (F (3,49) = 4.76, p = .005, R^2adj = .18). The analysis showed that both CRT score (Beta = .30, t (52) = 2.25, p = .029) and estimated accuracy (Beta = .29, t(52) = 2.2, p = .03) were significant predictors. However, the measure of sensitivity under random instructions did not predict sensitivity under logic instructions (Beta = .19, t (52) = 1.4, p = .15). The analysis on sensitivity under random instructions yielded a non-significant regression model (F (3,49) = 1.45, p = .24). These findings broadly confirm the earlier analyses based on a median split of estimated accuracy, although the regression analysis on the logic instruction condition does indicate a relatively small predictive relationship between estimated and observed accuracy.

Discussion

Experiment 1 was designed to test the 'intuitive logic' hypothesis using a novel instructional method that required participants to randomly respond to a series of simple syllogistic reasoning problems. The data establishes that, with simple inferences under logic instruction, people are able to discriminate between valid and invalid arguments well. Beliefs also have an impact on their judgments, but for simple one model syllogisms the effect is relatively small. In line with Wegner's research we also found that logical validity, and to a smaller extent conclusion believability, impacted on endorsement rates when participants were instructed to generate a random sequence of responses. Signal detection analysis confirmed that sensitivity to logic under random instruction was a reliable and robust finding. The results suggest that with these simple inferences, the underlying logical structure is processed at an intuitive level, or to adopt Wegner's terminology: people appear to demonstrate a level of 'uncontrolled [logical] intelligence'. Seemingly, this occurs outside of conscious awareness (similar to Wegner's findings), confirmed by the analysis on the post block estimates, which shows that people seem to have little self-insight into the impact of validity on random responding.

It is well known that high capacity reasoners tend to perform better on reasoning tasks (Stanovich, 1999, 2009). Typically, those with higher IQ or with a greater disposition to be more reflective in their thinking style, are more inclined to override a defective heuristic response (Evans, 2007) and are more willing to engage in Type 2 – analytical processing (Toplak et al., 2011, 2014). In other words, the association between better reasoning and cognitive ability, is typically seen as a Type 2 process overriding a Type 1 process, characteristically based on defective beliefs and associations. However, recent research by Thompson et al. (2018) suggests that individual differences in intuitive logic are related to variations in cognitive capacity. They demonstrated, across syllogistic and base rate problem types, that for high capacity reasoners, logic interfered more with belief judgments and vice versa for low capacity reasoners. They argued that cognitive capacity may determine how readily a reasoner automatises logical processing allowing for logical responses to be generated by Type 1 processing. Although our study did not use the range of individual differences measures that Thompson et al. (2018) did, we did examine the relationship between the CRT, a commonly used measure of analytical cognitive style, and sensitivity. Whilst there was a significant relationship between the CRT and variations in sensitivity under logic instruction, there was no relationship under random instructions. Whilst the absence of a correlation is relatively weak evidence, the findings are at least consistent with the idea that the influence of logical structure under the two types of instruction arise from independent processes.

Experiment 2

Experiment 1 employed a novel random judgment task to support the claim that people are intuitively sensitive to logic. Our findings demonstrated that the logical validity of simple syllogistic arguments influences people's capacity to generate random responses when instructed to do so. The key objectives of Experiment 2, were to replicate the findings from the first experiment, and examine whether they would extend to more complex syllogistic arguments.

Trippas et al. (2017) demonstrated that argument complexity was a key component in determining the boundary conditions of logical intuition. They instructed participants to respond based upon the believability or logical validity of a presented conclusion and showed that on simple conditional inference problems (Modus Ponens), logic impacted on belief judgments more than belief impacted on logic judgments. As discussed earlier, this finding is consistent with the claim that the logic of an argument is processed early and intuitively. In contrast, on more complex syllogisms, beliefs impacted more on logic judgments than vice versa, the reverse of what was observed on simple arguments. This finding indicates that on complex problems the logical conclusion is not available early enough to interfere with belief judgments and hence is not generated rapidly and intuitively.

In Experiment 2 we assigned participants to two conditions of complexity, one as a replication of Experiment 1 with simple, single-model syllogisms, and a second condition using a set of complex mutli-model syllogisms. In both conditions, participants were presented with the stimuli under logic and random instruction in a blocked design. Trippas et al. showed that for more complex arguments, conflict between belief and logic interfered more with the validity of the argument under logic instruction. This finding was explained as arising because complex logical judgments require deliberative thinking, whereas belief responses are available

Table 3. Examples of the complex syllogisms used in Experiment 2.

Valid		Invalid	
Believable	Unbelievable	Believable	Unbelievable
Some rodents are dobber	Some mice are dobber	Some thieves are ziji	Some criminals are ziji
No dobber are mice	No dobber are rodents	No ziji are criminals	No ziji are thieves
Some rodents are not mice	Some mice are not rodents	Some criminals are not thieves	Some thieves are not criminals

more immediately. Accordingly, we predicted that under random instruction there would be no impact of logic on responses, whereas for simple problems we expect an impact of logical validity on random responding as observed in Experiment 1.

Method

Participants

A total of 100 undergraduates from Macquarie University took part in Experiment 2. Eight one females and 19 males (mean = 21 yrs.) were recruited through the online Psychology system SONA. Participants received 1 course credit for 30 minutes of their time.

Design, materials & procedure

A 2 (Validity: Valid/Invalid) x 2 (Believability: Believable/Unbelievable) x 2 (Instruction: Logic/Random) x 2 (Complexity: Simple/Complex) mixed design was used in Experiment 2, with repeated measures on the first three factors. We randomly assigned participants to the simple and complex argument conditions. In the simple condition, participants were presented with the same 64 simple syllogisms as used in Experiment 1. In the complex condition participants were presented 64 complex multi-model syllogisms (for example, see Table 3). As with Experiment 1, all stimuli consisted of half conflict problems and half no-conflict problems. Conflict problems included 8 Valid-Unbelievable and 8 Invalid-believable items, and the no-conflict problems consisted of 8 Valid-Believable and 8 Invalid-Unbelievable items. In both conditions, the stimuli were presented in 2 blocks under logic and random instructions (with order counterbalanced between participants) and each participant was presented with a unique list of 64 items.

The complex syllogisms comprised of 16 syllogistic structures as used by Trippas et al. (2017) where the structures controlled for figural bias (see supplementary material A and also refer to Johnson-Laird & Byrne, 1991). The simple syllogisms consisted of the same structures used in Experiment 1.

Each participant received the same post-block estimate questions, as Experiment 1: the memory test after the random blocks and the 7-item Cognitive Reflection Test (see Experiment 1 for instructions).

Results

A 2 (Validity: Valid/Invalid) x 2 (Believability: Believable/Unbelievable) x 2 (Complexity: Simple/Complex) mixed ANOVA was carried out on both Logic and Random instruction conditions. Order of instruction was counterbalanced and not included as a factor in the main analysis. However, as per experiment 1, where appropriate we report the magnitude of the interaction between validity and order. The analysis of endorsement rates and Log latencies are reported separately for each instructional condition.

Two participants were removed from the *simple* condition for scoring 2SD (-2.0) below the sample mean ($M = 82\%$) on the memory test, leaving a total of $n = 48$ in the simple condition and $n = 50$ in the complex condition. Table 4 presents the endorsement rates and untransformed latencies for each condition in Experiment 2.

Logic instruction

A 2 (Validity: Valid/Invalid) by 2 (Believability: Believable/Unbelievable) by 2(Complexity: Simple/Complex) mixed analysis of variance with repeated measures on the first two factors, revealed a main effect of *validity*; $F(1, 96) = 202.687$, MSE $= 737.6$, $p < .001$, $\eta_p^2 = .679$, where valid conclusions were endorsed more than invalid conclusions (75% vs. 36%) and a main effect of *believability*; $F(1, 96) = 25.231$, MSE $= 576.4$, $p < .001$, $\eta_p^2 = .208$, where believable conclusion were endorsed more than unbelievable conclusions (62% vs. 50%).

There was also a main effect of *complexity*; $F(2, 96) = 22.508$, MSE $=$ 567.6, $p < .001$, $\eta_p^2 = .190$, which showed that overall endorsement rates where significantly higher for complex inferences (62%) compared to the simple inferences (50%). *Complexity* also interacted with *validity*; $F(1, 96) = 78.719$, $p < .001$, $\eta_p^2 = .451$; and *believability*; $F(1, 96) = 7.965$, $p = .006$, $\eta_p^2 =$ $.077$. Pairwise comparisons taken from separate analyses carried out on the simple and complex conditions, confirmed that the *validity* effect was present for both ($p < .001$) but the mean difference between valid and invalid arguments was larger for *simple* inferences ($M_D = 63$) compared to the *complex* inferences ($M_D = 15$). The main effect of *believability* was also significant for both inference types; however, the effect was bigger with *complex* inferences ($M_D = 19$; $p < .001$) compared to the *simple* inferences ($M_D = 5$, $p =$ $.031$). These effects are aligned with previous findings which show reduced

Table 4. Endorsement rates and raw latencies, across validity and believability, under Logic and Random Instruction, for both inferences. SD in brackets.

	Simple				Complex			
Instructions	VB	VU	IB	IU	VB	VU	IB	IU
Logic – Endorsements (%)	**86(16.8)**	**77(21.1)**	**19(23.8)**	**17(19.6)**	**77(19.2)**	**61(30.5)**	**66(24.8)**	**43(23.7)**
Logic – Latencies(ms)	**10,149(6086)**	**9,880(4893)**	**10,427(7064)**	**10,375(5005)**	**11,682(8309)**	**11,218(7830)**	**14,063(11762)**	**13,783(9931)**
Random – Endorsements (%)	**61(23.3)**	**62(18.8)**	**45(20.6)**	**38(21.9)**	**54(15.3)**	**49(16.7)**	**56(16.0)**	**51(16.3)**
Random – Latencies(ms)	**5,782(4708)**	**5,509(4948)**	**5,745(5211)**	**6,140(6376)**	**4,630(7386)**	**4,207(5476)**	**4,229(7297)**	**4,740(7436)**

Note: VB = Valid-Believable, VU = Valid-Unbelievable, IB = Invalid-believable, IU = Invalid-Unbelievable.

belief bias effects for simple syllogisms compared to more complex syllogistic inferences (Johnson-Laird & Byrne, 1996; Newstead et al., 1992).

There was no interaction between *validity* and *believability*; $F(1, 96) = .005$, MSE = 201.464, $p = .943$, $\eta_p^2 < .001$. However, there was a three-way interaction with complexity; $F(1, 96) = 6.364$, $p = .013$, $\eta_p^2 = .062$. For *simple* inferences the effect of believability was larger for valid items (VB = 86 vs. VU = 77) compared to invalid items (IB = 19 vs. IU = 17). In contrast, for *complex* inferences the effect of belief was larger on invalid items (IB = 66 vs. IU = 43) compared to valid ones (VB = 77 vs. VU = 61). The interaction between *believability* and *validity* was significant for *simple* inferences; $F(1,47) = 4.155$, $p = .047$, $\eta_p^2 = .081$; but for *complex* inferences was non-significant; $F(1,49) = 2.685$, $p = .108$, $\eta_p^2 = .052$.

Analysis of the latency data revealed a main effect of *validity*; $F(1, 96) = 7.967$, MSE = .020, $p = .006$, $\eta_p^2 = .077$; with people taking longer to respond on invalid problems (M = 12,162ms) compared to valid problems (10,732ms). There was no main effect of *complexity*; ($F < 1$) but there was a marginal interaction between *complexity* and *validity*; $F(1, 96) = 3.282$, MSE = .020, $p = .073$, $\eta_p^2 = .033$. A separate analysis on *each condition* confirmed, consistent with Experiment 1, that the effect of *validity* was not significant for *simple* inferences; $F(1, 47) = .392$, MSE = .026, $p = .534$, $\eta_p^2 = .008$, but was reliable for *complex* inferences; $F(1, 49) = 14.399$, MSE = .015, $p < .001$, $\eta_p^2 = .237$. There were no other main effects or interactions to report ($F < 1$) .

Random instruction

The 2 (Validity: Valid/Invalid) by 2 (Believability: Believable/Unbelievable) by 2(Complexity: Simple/Complex) mixed analysis of variance on endorsement rates under random instructions revealed a main effect of *validity*; $F(1, 96) = 12.571$, MSE = 662.2, $p = .001$, $\eta_p^2 = .116$, showing that overall participants endorsed more valid arguments (M = 57%) than invalid arguments (M = 48%). There was a main effect of *believability;* $F(1, 96) = 6.413$, MSE =275.9, $p = .013$, $\eta_p^2 = .063$, confirming that believable arguments were endorsed more than unbelievable arguments (54% vs. 50%). There was no main effect of *complexity*; $F(1, 96) = .550$, MSE =182.0, $p = .460$, $\eta_p^2 = .006$; however, *complexity* did interact with *validity*; $F(1, 96) = 18.205$, $p < .001$, $\eta_p^2 = .159$. Follow up analyses confirmed that the validity effect was significant for *simple* inferences; $F(1, 47) = 18.486$, MSE =1071.3, $p < .001$, $\eta_p^2 = .282$, replicating the findings from Experiment 1, but was non-significant for *complex* inferences ($F < 1$; see Figure 1). There was no interaction between *validity* and *believability;* $F(1, 96) = 1.227$, MSE =287.1, $p = .271$, $\eta_p^2 = .013$, and *complexity* did not interact with *believability*; $F(1, 96) = .449$, $p = .504$, $\eta_p^2 = .005$. The effect of *validity* on random responding for *simple* inferences,

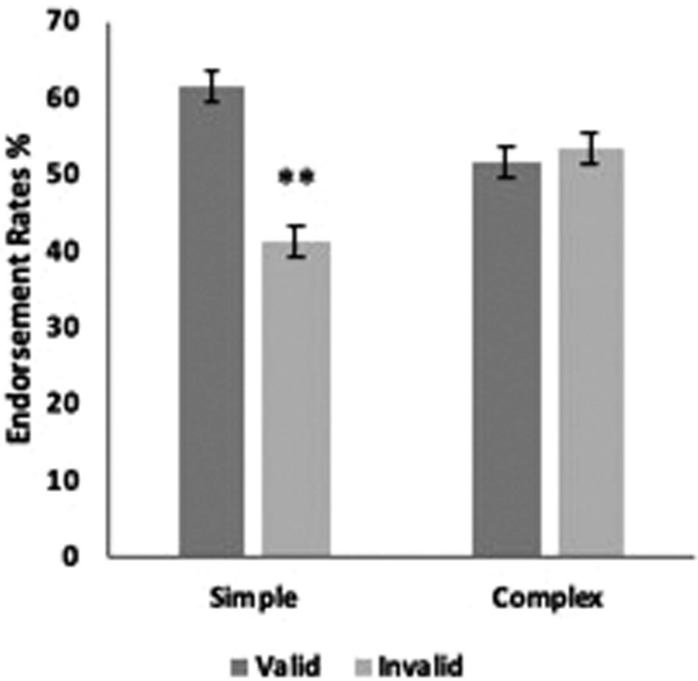

Figure 1. Conclusion endorsement rates (%) under random instructions for simple and complex arguments as a function of validity.

like Experiment 1, extended across the majority of participants, with 79% of participants showing logical accuracy rates above 50%. In contrast only 54% of participants responded in line with the believability of the conclusion more than 50% of the time.

As per experiment 1, we wanted to confirm that completing a logic instruction block prior to the random instruction block could not explain the validity effect under random instruction. Re-running the analysis, including order, on simple inferences alone showed a marginal interaction between order and validity; ($F(1,46) = 3.26$, MSE = 1022, $p = .08$, $\eta_p^2 = .066$). However, as in Experiment 1, this reflected a somewhat larger difference in endorsements between valid and invalid items when the random instruction condition occurred first ($M_v = 65$, $M_{inv} = 36$) than when it followed a logic instruction block ($M_v = 58$, $M_{inv} = 46$). Thus, there is no evidence that the effect of validity on random judgments arises because completing logical judgments cues subsequent structural processing under random instructions.

An analysis of response latencies under random instructions showed no main effect of *validity;* ($F < 1$), no effect of *believability* ($F < 1$) or *complexity*; $F(2, 96) = 1.578$, MSE = .695, $p = .212$, $\eta_p^2 = .016$; and there were no interactions to report (all $F < 1$).

Once again it is worth noting that latency of response under random instruction (M_{com} = 4,452ms, M_{sim} = 5,794ms) was significantly lower than logic instructions for both *complex* (M_{com} = 12,687ms; t(49) = 10.66, *p* <.001) and *simple* inferences (M_{sim} = 10,208; t(47) = 5.729, *p* <.001) indicating that participants spend substantially less time responding under random compared to logic instructions.

Task performance, self-insight and cognitive style

As in Experiment 1 our second set of analyses are designed to evaluate the extent to which participants demonstrate insight into their responding. Before considering the accuracy of performance estimates, we provide confirmation of the key findings from the ANOVA analyses, drawing on the measures of sensitivity and bias discussed earlier. Under random instructions sensitivity was significantly above zero for *simple* inferences (M =.64, t(47) = 4.02, *p* <.001) but not for *complex* inferences (M = −.04, t(49) = −.67, p = .508). Under logic instructions sensitivity was significantly above zero for both *simple* (M = 2.14, t(47) = 12.18, *p* < .001) and *complex* inferences (M = .46, t(49) = 4.97, *p* < .001).

In contrast for both *simple* and *complex* inferences under random instructions, there was no significant difference in bias between believable (M_s = −.08; M_c =-.14) and unbelievable (M_s = .017; M_c =-.004) arguments (*Simple*: t(47) = −1.705, *p* = .095; *Complex*: t(49) = −1.854, *p* =.07), although these effects would be significant under a one-tailed test. Under logic instructions there was a significant two tailed effect of beliefs for both *simple* (M_{sb} = −.09, M_{su} = .08, t(47) = −2.421, *p* = .019, and *complex* inferences (M_{cb} = −.63, M_{cu} = −.05, t(49) = −4.581, *p* < .001.

Our next set of analyses examined whether participants showed self-insight into their performance by examining the post-task judgments of accuracy under logic and random instructions, and judgments of randomness under random instructions. Since there was no reliable effect of logic under random instruction for *complex* inferences, we focus the analyses in the following section on *simple* inferences only.

Comparable with Experiment 1, observed accuracy under random instructions (M = 60) was significantly higher than estimated accuracy (M = 39, t(48) = 8.59, *p* < .001), which was also the case under logic instructions (M_o = 82, M_e = 57, t(48) = 9.24, *p* < .001). Again, there was no evidence that the difference between observed and estimated accuracy was any greater under random than logic instructions (M_{dr} = 21, M_{dl} = 25, t = 1.28, p = .209).

A median split on estimated accuracy scores was carried out and sensitivity was compared for those that judged high accuracy (High group) compared to the Low judgement group (Low group). This would help establish whether participants had insight into their level of logical accuracy. Under logic instructions the mean accuracy rating for the High group was 72 ($N = 26$) and the Low group ($N = 22$) was 39, with a significant difference in sensitivity between the groups ($M_H = 2.73$, $M_L = 1.44$, $t(46) = 4.276$, $p < .001$). Under random instructions, the High ($N = 21$, $M_e = 54$) and Low groups ($N = 27$, $M_e = 28$) showed no reliable difference in sensitivity ($M_H = .94$, $M_L = .42$, $t(46) = 1.65$, $p = .106$), although there was a trend to greater sensitivity in the High group.

Examining the random estimates on the random trials, revealed a significant difference in sensitivity between the groups ($M_H = .266$, $M_L = 1.05$, $t(46) = -2.603$, $p = .012$). With the High group ($N = 21$, $M_e = 54$) showing less sensitivity to logical validity compared to the Low group ($N = 27$, $M_e = 28$).

Overall people explicitly underestimated the accuracy of their performance under both logic and random instruction. However, underestimation doesn't indicate a complete lack of insight, and the association between estimates of performance and levels of sensitivity suggest that, in this instance, participants possess some awareness of the impact of validity on their observed accuracy. Whilst this did not reliably extend to accuracy under random instruction, interestingly, there does appear to be some level of indirect insight into the impact of validity on levels of randomness. With those estimating low randomness scores showing higher levels of sensitivity to logic.

Turning now to the relationship between the CRT and sensitivity, there was a significant correlation between performance on the CRT and sensitivity under logic instructions ($r(46) =. 443$, $p = .002$), but no reliable correlation under random instruction ($r(46) = -.088$, $p = .55$). These correlation coefficients were significantly different as shown by a Fishers R to Z transformation ($z = 2.68$, $p = .004$). This corroborates the findings from Experiment 1, suggesting that the two tasks are measuring different modes of logical sensitivity. Again, this claim that is consistent with zero order correlation between the measures of sensitivity across the two instructional tasks ($r = .17$, $p = .237$). Importantly, the sensitivity indices showed respectable levels of split half reliability under logic (*alpha* = .788) and random (*alpha*= .809) instructions.

Finally, as in Experiment 1, we explored the extent to which logical sensitivity under the different instructional conditions could be predicted through a combination of the measures of estimated accuracy, CRT score and sensitivity. The multiple regression analysis on the logic instruction condition revealed a significant model ($F(3,44) = 12.87$, $p < .001$, $R^2adj = .43$). As in Experiment 1, CRT score ($Beta = .35$, $t(47) = 3.11$, $p = .003$) and estimated accuracy ($Beta = .49$, $t(47) = 4.3$, $p < .001$) were significant predictors. In contrast to Experiment 1, sensitivity under random instructions showed a small but reliable predictive relationship (Beta = .27, t (47) = 2.4, p = .02).

Under random instructions the analysis also resulted in a significant regression model (F (3,47) = 5.16, p = .004, R^2adj = .21). The only significant predictor was estimated accuracy (Beta = .46, t (47) = 3.41, p = .001). Neither CRT score (t (47) = −.48, p = .64) nor sensitivity under logic instructions (t (47) = 1.03, p = .31) were reliable. The results of the regression analyses confirm some of the findings of experiment 1, notably that accuracy estimates and CRT scores are predictive of sensitivity under logic instructions. However, in contrast to the previous study, under random instructions, there is good evidence that accuracy estimates predict sensitivity, suggesting that participants have some awareness of the impact of the logical structure on their random responding.

Discussion

The primary aim of Experiment 2 was to establish whether the logic effect observed under random instructions in Experiment 1, was a robust and replicable finding. The secondary aim was to determine if these findings would extend to more complex problem types. Under logic instruction the results confirmed that people readily discriminate between valid and invalid arguments for both simple and complex problems, although the effect is larger with simple syllogisms. In contrast, the effect of beliefs was larger for complex syllogisms, which is consistent with previous literature showing bigger belief bias effects with more complex syllogisms (Johnson-Laird & Byrne, 1996; Newstead et al., 1992).

The effect of logical validity on random responding replicates the findings of Experiment 1, confirming that on simple problems, participants' capacity to generate a random sequence of responses is influenced by logical structure. However, as expected, the effect was only significant for simple syllogisms suggesting, in line with earlier research (Trippas et al., 2017), that the logical conclusion on complex arguments is not generated intuitively.

In Experiment 2, on simple arguments and in contrast to some of the findings in Experiment 1, participants appear to show insight into their performance. Although performance accuracy was significantly underestimated, there was a relationship between judged and observed accuracy. This relationship was also present under random instructions as shown in the regression analysis and there was also evidence of some indirect insight into the influence of logical structure on performance, as the group of participants who judged themselves to be less random displayed higher levels of logical sensitivity. These findings suggest that participants are to some extent able to detect and report the influence of logical problem features on their responding.

Experiment 2 also confirmed a key finding from Experiment 1; the presence of a significant relationship between the CRT and sensitivity under logic instructions, in the absence of a relationship under random instructions. This is consistent with the idea that there is, at the very least, a partial dissociation between the processes underlying responding on the two tasks and supports the idea that sensitivity to logical structure under each instructional condition may reflect the output of distinct processing systems.

General discussion

The main objective of the research reported in this paper was to evaluate evidence for an intuitive effect of logical validity on a novel task in which participants were explicitly instructed to respond randomly. This differs from previous studies where participants are often instructed to respond logically and evidence for intuitive logic is claimed from the observation of increase latencies (De Neys, 2012; De Neys & Glumicic, 2008) or reduced confidence (De Neys et al., 2011) on conflict judgments even where a heuristic response is given. Such responses have been presumed to reflect the parallel computation of a logical response that is insufficiently salient to guide responding (De Neys, 2012, 2014; Handley & Trippas, 2015; Pennycook et al., 2014), but increases uncertainty and response times. Researchers have also claimed evidence for intuitive logic from the observation that logical validity influences responding despite the presence of high working memory load or restricted time (Franssens & De Neys, 2009; Bago & De Neys, 2017; Bago & De Neys, 2019; De Neys, 2017) where presumably the opportunity to engage in deliberative thinking is limited. Whilst this evidence is persuasive it is possible that participants, rather than engaging in intuitive processing, are, in fact, reasoning deliberatively but doing so rapidly.

In the studies reported here we adapted the random response paradigm, from Wegner et al. (2003), and showed that participants were more likely to endorse valid compared to invalid conclusions despite being instructed to generate a random sequence of responses. This effect was consistent for simple arguments across Experiments 1 and 2 and provides strong confirmation of the 'intuitive logic' hypothesis using a task that does not instruct participants to respond logically. Interestingly, when the complexity of the argument was increased (Experiment 2) the intuitive logic effect, under random instruction, was no longer evident. This implies that sensitivity to logical structure is either unavailable or not strong enough to interfere with random responding, when the problem features are complex. These findings are consistent with previous work, showing that intuitive logic effects are only discernible with simple inferences and not complex problem types (Trippas et al., 2017).

The overall latency of responding under the two instructional conditions is consistent with an absence of deliberative judgment under random instructions, with logic judgments taking more than double the time to complete compared to random judgments. However, there was some evidence of insight into accuracy of performance when responding randomly, as shown through the relationship between estimated accuracy and sensitivity in Experiment 2. This suggests that participants had at least some sense of the influence that logical structure was having on their responding. There was also evidence of indirect insight into the effect logic was having on the randomness of people's judgments in Experiment 2, with those people estimating low levels of randomness displaying greater sensitivity to the validity of the argument. This is similar to the findings using the conflict detection paradigm, where reasoners appear to be implicitly aware of conflict, through the reports of lower confidence ratings (De Neys et al., 2011) or longer latencies (De Neys & Glumicic, 2008); indicating a sensitivity to logical structure, at an intuitive level. One possibility is that internal cues, such as increased uncertainty, or reduced confidence support an implicit awareness that there is a competing response (i.e. logical validity) that is then contributing to subsequent judgments about accuracy or randomness of judgments.

In support of a dissociation between intuitive and explicit logic, there was no relationship between the measures of sensitivity for random and logic instructions despite both sensitivity measures displaying good internal reliability. This claim is consistent with the evidence of a relationship between sensitivity under logic instructions and individual differences on the CRT in both Experiments 1 and 2, a finding that was absent for random instructions. This shows that accuracy under logic instructions varies as a function of analytic cognitive style, a measure reflecting individual variations in engagement with deliberative thinking.

One important question to consider, is whether the findings we report here reflect a genuine effect of intuitive logic. For example, could participants have been primed to process the logical validity of the arguments under random instructions because they also received similar problems under logical instructions? The evidence suggests not, because the logic effect was stronger, in both experiments, when participants completed the random block first, which is the opposite of what one might expect if the logical instruction was acting as a prime during the random block of problems. Is it possible that the effects arise because a significant minority of participants are explicitly engaged in reasoning under random instructions, whilst the remainder are responding randomly? Once again the findings suggest not, as the validity effect extended across the majority of participants in both experiments, with 63% showing greater than 50% accuracy rates in Experiment 1 which rose to 79% in Experiment 2.

Finally, could it be that people weren't following the instructions? In both experiments, participants were explicitly instructed to respond randomly, with details directing them to be as free and random as possible, in the generation of a random sequence of responses. We know that participants successfully followed the logical instructions, generating a strong logic effect as well as the typically observed effects of complexity and beliefs on responding for simple and complex syllogistic inferences and there was some level of direct insight (Experiment 2) into the accuracy of their performance on simple inferences. There is no reason to think that participants, whilst able to follow a complex set of logical instructions, were then unable to engage with a much simpler direction to generate a random sequence of responses.

Using the same paradigm Wegner et al., similarly found that people's general knowledge impacted on their random responding and claimed it did so outside of their conscious awareness. Similarly, 50 years of research on heuristics and unconscious biases (Tversky & Kahneman, 1974; Kahneman, 2011) recognises the prevalence of knowledge impacting on judgment outside of awareness, so much so the research has been used to inform policy, change health practices and influence consumer behaviour (see Thaler & Sunstein, 2009). Hence there is a very clear precedent for problem features to influence judgment intuitively and we would argue that, in the case of simple arguments, the logical structure is so salient that it operates to influence judgment in a similar way.

Interestingly, in our experiments the most significant impact on random responding, was related to the logical structure of the arguments presented, whereas beliefs only had a small and variable impact. Traditional Dual Process models would suggest that beliefs should have the biggest intuitive influence, since they are often the more salient cue to responding.

So, why didn't beliefs have as big an impact on random judgments, as they did in Wegner et al's., study? One possibility concerns the way in which the questions are presented. For example, in our experiments we presented participants with logical statements that consisted of premises and conclusions, in the form of determinately valid or invalid syllogisms. Wegner on the other hand, presented participants with one sentence statements such as: 'Does a triangle have 3 sides?'. In essence this would be similar to presenting participants with only the conclusion of the syllogism, removing the potential for simple structural processing to interfere with the influence that beliefs may have on their responding. Perhaps the inference is cued intuitively following the presentation of the first two premises and hence the validity of the conclusion is processed prior to the participant being able to process the believability of the conclusion. Thus the believability of the conclusion has limited impact on random responding. We predict that if we were to run a similar experiment, where the conclusion alone was presented, we would replicate Wegner's findings with sentence believability influencing random responding.

So, what implications do our findings have for Dual Process Theories? First, the results challenge traditional Default Interventionist models that emphasise Type 1 processing as typically relying on learned associations and relevant beliefs. Instead, they corroborate recent claims for the existence of intuitive logic accomplished at a Type 1 level of processing (Bago & De Neys, 2017; 2019). The findings also provide important confirmation of one of the boundary conditions of this phenomenon, that of argument complexity, suggesting that at a certain level of complexity, the processes involved in generating a logical conclusion cannot be accomplished intuitively (see, Brisson et al., 2018; Trippas et al., 2017).

For simple logical arguments the findings suggest that people may possess a level of 'uncontrolled [logical] intelligence' which can impact on an individual's capacity to respond randomly, even when explicitly instructed to do so. However, this intuitive sensitivity to logical structure, is only evident when the logical structure is simple. We would argue that an intuitive response is generated in line with the most salient cue and on simple arguments the logical cue may be stronger than one based upon beliefs, due to a higher degree of saliency or fluency (Bago & De Neys, 2017). This then gives rise to instances of intuitive logic; but if the heuristic cue is stronger, as it will be for more complex logical arguments, then this will typically generate the classic belief-bias effect. Over the last few years, researchers have suggested the move towards 'hybrid' DP models of explanation, which are typically a blend of serial and parallel processing that can account for instances of logical intuitions (Banks & Hope, 2014; De Neys, 2012; Handley et al., 2011; Handley & Trippas, 2015; Howarth et al., 2019; Pennycook et al., 2012; Thompson et al., 2018). Typically, the models suggest there are multiple Type 1 outputs (De Neys, 2012; Newman et al., 2017) which can be based on heuristics (beliefs/associations) and logic (basic logical principles and probabilities) as well as various other intuitive cues (i.e. visual/kinaesthetic information). These multiple intuitive cues, triggered simultaneously, can create conflict, if they support incongruent responses. Deliberative Type 2 processing is required to deliver an explicit response, and this may involve rationalising an intuitive response or inhibiting and over-riding the intuitive cue in favour of additional deliberation.

According to Handley et al. (2011, 2015) parallel competitive DP model, complexity is a key feature in determining the direction of interference from an intuitive cue. Research using the instructional paradigm (Trippas et al., 2017) has shown that with simple arguments, logic produces more interferences than beliefs, whereas with complex arguments the opposite is true. Our current findings confirm that logic creates more interference on randomness judgments when the problems are simple (Experiment 1 & 2), an effect that disappears when the complexity or the inference is increased (Experiment 2). The principle aim of this paper was to demonstrate that logical structure could influence a person's ability to respond randomly and that, that impact is mitigated by argument complexity. The experiments were not designed to differentiate between different types of DP models. However, the findings are consistent with DP accounts that recognise the potential, intuitive influence that logical structure can have on people's judgments (Bago & De Neys, 2017; De Neys, 2012; De Neys and Pennycook, 2019; Pennycook et al., 2014; Pennycook, 2017; Thompson et al., 2011). However, Handley et al. (2011, 2015) parallel competitive DP model, highlights the importance of problem complexity as a key feature in identifying the more salient cue provided by the problems presented. In respect to our current finding, the data fits well with a model that highlights the importance of problem complexity on the impact of logical intuitions.

The pattern of findings in these experiments are consistent with the idea that people are sensitive to the logical validity of simple but not complex arguments. However, an important remaining question concerns the nature of the cognitive mechanisms that underlie 'intuitive logic'. In our earlier discussion on logical intuition, we referenced the foundational work of Braine and O'Brien (1991) and Rips (1994), whose models of natural deduction are built around the proposal that certain inferences are accomplished through the application of 'direct' rules of inference that are activated automatically in the course of processing the meaning of logical connectives. These accounts distinguish between simple inference rules and more complex 'indirect' rules that require more deliberative strategic processing. One could argue that the evidence of validity effects for simple, but not complex problems under random responding is supportive of this distinction. However, whilst there is increasing evidence, including the findings we report here, that reasoners appear to show sensitivity to logical structure, we think it is premature to conclude that these findings reflect the activation of rules that align with the normative standards of propositional logic. Our complex and simple problems differ in a number of ways and it is possible that reasoners are sensitive to problem features that happen to align with logical validity, but do not reflect normative rules.

Consider for example, the simple arguments shown in Table 1. The conclusion to the valid argument can be shown to be necessary by elementary natural deductions rules, but is also consistent with simple heuristic accounts of syllogistic reasoning such as atmosphere or the min-heuristic (see, for example, Chater & Oaksford, 1999). However, the conclusion to the invalid argument is impossible given the premises and is misaligned with the type of conclusion supported by a simple heuristic strategy. In contrast, the more complex arguments have invalid conclusions that are consistent with the premises, but not necessitated by them. Such conclusions are also consistent with the atmosphere and min-heuristics.

The simple and complex arguments differ in another important way; on simple arguments, the valid conclusion is logically possible and the invalid conclusion impossible given the premises. In contrast, on complex argument forms, the valid and invalid conclusions are both logically possible given the premises. Hence the presence of a validity effects under random responding for simple problems may arise for reasons other than sensitivity to logic. Instead, simple heuristics may be automatically activated and lead reasoners to select the valid conclusion more often than the invalid one. Alternatively, perhaps the effect arises because the valid conclusion to the simple problems is possible and hence consistent with the premises in contrast to the invalid conclusion, which is not. Future research could usefully test these alternative accounts by manipulating the alignment of conclusions with simple heuristics, or through comparing 'intuitive logic' effects for conclusions that are necessary compared to those that are just possible.

Conclusion

There continues to be significant debate in the literature concerning the existence of 'logical intuition' and the extent to which such intuitions are delivered independently of the processes involved in deliberative thinking. The emphasis on distinct processing systems aligns this issue with a broader debate concerning the role of dual processes in human thinking. The present studies provide additional evidence for the existence of logical intuitions using a unique approach that does not share the limitations of previous research where participants are explicitly instructed to reason logically. Our findings might suggest a more optimistic view of human reasoning; our participants cannot avoid being logical, even though such apparently intelligent behaviour may be somewhat uncontrolled.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Stephanie Howarth http://orcid.org/0000-0001-6149-0064

References

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

Bago, B., & De Neys, W. (2019). The smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*(3), 257–299. https://doi.org/10.1080/13546783.2018.1507949

Bago, B., Frey, D., Vidal, J., Houdé, O., Borst, G., & De Neys, W. (2018). Fast and slow thinking: Electrophysiological evidence for early conflict sensitivity. *Neuropsychologia*, *117*, 483–490. https://doi.org/10.1016/j.neuropsychologia.2018.07.017

Banks, A. P., & Hope, C. (2014). Heuristic and analytic processes in reasoning: An event-related potential study of belief bias. *Psychophysiology*, *51*(3), 290–297. https://doi.org/10.1111/psyp.12169

Barrouillet, P. (2011). Dual-process theories and cognitive development: Advances and challenges. *Developmental Review*, *31*(2-3), 79–85. https://doi.org/10.1016/j.dr.2011.07.002

Braine, M. D., & O'Brien, D. P. (1991). A theory of if: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*(2), 182–203. https://doi.org/10.1037/0033-295X.98.2.182

Brisson, J., Schaeken, W., Markovits, H., & De Neys, W. (2018). Conflict detection and logical complexity. *Psychologica Belgica*, *58*(1), 318–322. https://doi.org/10.5334/pb.448

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, *38*(2), 191–258. https://doi.org/10.1006/cogp.1998.0696

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, *62*, 271–298. https://doi.org/10.1146/annurev.psych.121208.131622

De Neys, W., & Pennycook, G. (2019). Logic, Fast and Slow: Advances in Dual-Process Theorizing. *Current Directions in Psychological Science*, *28*(5), 503–509. https://doi.org/10.1177/0963721419855658

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner . *Psychological Science*, *17*(5), 428–433. https://doi.org/10.1111/j.1467-9280.2006.01723.x

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*(1), 28–38. https://doi.org/10.1177/1745691611429354

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

De Neys, W. (2017). Bias, conflict, and fast logic: Towards a hybrid dual process future? In *W. Neys (Ed.),* Dual process theory 2.0 (pp. 47–65). Routledge.

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PLoS One*, *6*(1), e15954. https://doi.org/10.1371/journal.pone.0015954

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3), 1248–1299. https://doi.org/10.1016/j.cognition.2007.06.002

De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 208–216. https://doi.org/10.3758/CABN.10.2.208

De Neys, W. D., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science*, *19*(5), 483–489. https://doi.org/10.1111/j.1467-9280.2008.02113.x

Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: it's a response bias effect. *Psychological Review*, *117*(3), 831–863. https://doi.org/10.1037/a0019634 20658855

Evans, J. S., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, *6*(3), 263–285. https://doi.org/10.1080/09541449408520148

Evans, J. S. B. T., Handley, S. J., Neilens, H., & Over, D. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *Quarterly Journal of Experimental Psychology (2006)*, *63*(5), 892–909. https://doi.org/10.1080/17470210903111821 19728225

Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Evans, J. S. B. T. (2006). Dual system theories of cognition: Some issues. *Paper presented at the the 28th Annual Meeting of the Cognitive Science Society* (pp. 202–207).

Evans, J. S. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Psychology Press.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

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

Evans, J. S. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, *11*(4), 382–389. https://doi.org/10.1080/13546780542000005

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

Franssens, S., & De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking & Reasoning*, *15*(2), 105–128. https://doi.org/10.1080/13546780802711185

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*(4), 237–241. https://doi.org/10.1111/j.1467-9280.1996.tb00366.x

Gangemi, A., Bourgeois-Gironde, S., & Mancini, F. (2015). Feelings of error in reasoning–in search of a phenomenon. *Thinking & Reasoning*, *21*(4), 383–396. https://doi.org/10.1080/13546783.2014.980755

Goel, V. (2003). Evidence for dual neural pathways for syllogistic reasoning. *Psychologica*, *32*, 301–309.

Goel, V., & Vartanian, O. (2011). Negative emotions can attenuate the influence of beliefs on logical reasoning. *Cognition & Emotion*, *25*(1), 121–131. https://doi.org/10.1080/02699931003593942

Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.

Handley, S. J., Newstead, S. E., & Trippas, D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(1), 28–43. https://doi.org/10.1037/a0021098

Handley, S. J., & Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. In Brian H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 62, pp. 33–58). Elsevier.

Hayes, B. K., Wei, P., Dunn, J. C., & Stephens, R. G. (2020). Why is logic so likeable? A single-process account of argument evaluation with logic and liking judgments. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *46*(4), 699–719. https://doi.org/10.1037/xlm0000753

Heit, E., & Rotello, C. M. (2014). Traditional difference-score analyses of reasoning are flawed. *Cognition*, *131*(1), 75–91. https://doi.org/10.1016/j.cognition.2013.12.003 24462712

Howarth, S., Handley, S. J., & Walsh, C. (2016). The logic-bias effect: The role of effortful processing in the resolution of belief-logic conflict. *Memory & Cognition*, *44*(2), 330–349. https://doi.org/10.3758/s13421-015-0555-x

Howarth, S., Handley, S., & Walsh, C. (2019). The logic sense: Exploring the role of executive functioning in belief and logic-based judgments. *Thinking & Reasoning*, *25*(4), 416–448. https://doi.org/10.1080/13546783.2018.1523808

Johnson-Laird, P., & Byrne, R. M. (1996). Mental models and syllogisms. *Behavioral and Brain Sciences*, *19*(03), 543–546. https://doi.org/10.1017/S0140525X00082091

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

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Morsanyi, K., & Handley, S. J. (2012). Logic feels so good-I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(3), 596–616. https://doi.org/10.1037/a0026099

Nakamura, H., & Kawaguchi, J. (2016). People like logical truth: Testing the intuitive detection of logical value in basic propositions. *PLoS One*, *11*(12), e0169166. https://doi.org/10.1371/journal.pone.0169166

Newman, I. R., 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*(7), 1154–1170. https://doi.org/10.1037/xlm0000372

Newstead, S. E., Pollard, P., Evans, J. S. B., & Allen, J. L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, *45*(3), 257–284. https://doi.org/10.1016/0010-0277(92)90019-E

Pennycook, G. (2017). A perspective on the theoretical foundation of dual-process models. *Dual Process Theory*, *2*, 34.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning? *Cognition*, *124*(1), 101–106. https://doi.org/10.1016/j.cognition.2012.04.004

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*(2), 544–554. https://doi.org/10.1037/a0034887

Polito, V., Barnier, A. J., & Connors, M. H. (2018). Hypnotic clever hands: Agency and automatic responding. *Journal of Experimental Psychology. General*, *147*(6), 815–828. https://doi.org/10.1037/xge0000451

Raoelison, M., & De Neys, W. (2019). Do we de-bias ourselves?: The impact of repeated presentation on the bat-and-ball problem.

Raoelison, M. T., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*, 104381. https://doi.org/10.1016/j.cognition.2020.104381

Reber, A. S. (1996). *Implicit learning and tacit knowledge: An essay on the cognitive unconcious*. Oxford University Press.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. MIT Press.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131. https://doi.org/10.1207/S15327957PSPR0402_01

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Lawrence Erlbaum Associates Publishers.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. In J. S. B. Evans & K. Frankis (Eds.), *In Two minds: Dual processes and beyond* (pp. 55–88). Oxford University Press.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, *24*(4), 423–444. https://doi.org/10.1080/13546783.2018.1459314

Stanovich, K. E., & Toplak, M. E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind & Society*, *11*(1), 3–13. https://doi.org/10.1007/s11299-011-0093-6

Stupple, E. J., & Ball, L. J. (2008). Belief–logic conflict resolution in syllogistic reasoning: Inspection-time evidence for a parallel-process model. *Thinking & Reasoning*, *14*(2), 168–181. https://doi.org/10.1080/13546780701739782

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thompson, V., & Morsanyi, K. (2012). Analytic thinking: Do you feel like it? *Mind & Society*, *11*(1), 93–105. https://doi.org/10.1007/s11299-012-0100-6

Thompson, V. A., Pennycook, G., Trippas, D., & Evans, J. S. B. (2018). Do smart people have better intuitions? *Journal of Experimental Psychology: General*, *147*(7), 945–961. https://doi.org/10.1037/xge0000457

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

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, *20*(2), 147–168. https://doi.org/10.1080/13546783.2013.844729

Topolinski, S., & Strack, F. (2008). Where there's a will – there's no intuition. The unintentional basis of semantic coherence judgments. *Journal of Memory and Language*, *58*(4), 1032–1048. https://doi.org/10.1016/j.jml.2008.01.002

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General*, *138*(1), 39–63. https://doi.org/10.1037/a0014678

Trippas, D., Kellen, D., Singmann, H., Pennycook, G., Koehler, D. J., Fugelsang, J. A., & Dube, C. (2018). Characterizing belief bias in syllogistic reasoning. *Psychonomic Bulletin and Review*, *25*, 2141–2174.

Trippas, D., Handley, S. J., & Verde, M. F. (2013). The SDT model of belief bias: Complexity, time, and cognitive ability mediate the effects of believability. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(5), 1393–1402. https://doi.org/10.1037/a0032398

Trippas, D., Handley, S. J., Verde, M. F., & Morsanyi, K. (2016). Logic brightens my day: Evidence for implicit sensitivity to logical validity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(9), 1448–1457. https://doi.org/10.1037/xlm0000248

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*(4), 539–552. https://doi.org/10.3758/s13421-016-0680-1

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Wegner, D. M., Fuller, V. A., & Sparrow, B. (2003). Clever hands: Uncontrolled intelligence in facilitated communication. *Journal of Personality and Social Psychology*, *85*(1), 5–19. https://doi.org/10.1037/0022-3514.85.1.5

Whittlesea, B. W. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(6), 1235.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*(6), 989–1000.