WORKING PAPERS

N° 147

Eye movements, pupil dilation, and conflict detection in reasoning: Exploring the evidence for intuitive logic

Zoe Purcell, Andrew J. Roberts, Simon J. Handley, Stephanie Howarth

IAST
INSTITUTE for ADVANCED STUDY in TOULOUSE

Eye movements, pupil dilation, and conflict detection in reasoning: Exploring the evidence for intuitive logic

Zoe Purcell[1,2], Andrew J. Roberts[3], Simon J. Handley[4], Stephanie Howarth[4]

[1] Institute for Advanced Study in Toulouse, University of Toulouse, Toulouse, France.

[2]Artificial and Natural Intelligence Institute of Toulouse, University of Toulouse, Toulouse, France.

[3] Department of Philosophy, Macquarie University, Sydney, Australia.

[4] School of Psychological Sciences, Macquarie University, Sydney, Australia.

Author notes. All authors contributed equally to this paper. Correspondence should be directed to stephanie.howarth@mq.edu.au.

Key words: Conflict Detection, Reasoning, Eye Tracking, Pupil Dilation, Dual Processes, Intuitive Logic.

Abstract

A controversial claim in recent dual process accounts of reasoning is that intuitive processes not only lead to bias, but are also sensitive to the logical status of an argument. The intuitive logic hypothesis draws upon evidence that reasoners take longer and are less confident on belief-logic conflict problems, irrespective of whether they give the correct logical response. In this paper we examine conflict detection under conditions in which participants are asked to either judge the logical validity or believability of a presented conclusion, accompanied by measures of eye movement and pupil dilation. The findings show an effect of conflict, under both types of instruction, on accuracy, latency, gaze shifts and pupil dilation. Importantly these effects extend to conflict trials in which participants give a belief-based response (incorrectly under logic instructions, or correctly under belief instructions) demonstrating both behavioural and physiological evidence in support of the logical intuition hypothesis.

1. Introduction

The idea that many of the judgments that we make are intuitive has a long history in psychology (e.g., Evans, 2008; Gilovich et al., 2002). Such intuitive judgments are claimed to play a role in social impressions (Lieberman, 2000), stereotypes (Greenwald & Banaji, 1995), learning (Reber, 1989), memory (Morwedge & Kahneman, 2010), and judgment and decision making (Kahneman & Frederick, 2005), together with many other cognitive, social, and perceptual processes. A common assumption underlying models of intuition is that judgments come to mind rapidly and automatically, and people are unaware of the origins of these thoughts (e.g., De Neys, 2012; Stanovich, 2018). In reasoning and decision-making research, it has been consistently shown that our intuitions can lead to systematic errors and biases in judgment. Such intuitions are automatic, come to mind effortlessly, and are often based upon simplifying heuristics that draw upon cues that do not have explicit justification or validity in the context of the task at hand (Kahneman, 2011). For example, in reasoning it is well known that the believability of a conclusion can influence its acceptance rate independently of its actual logical status (Evans et al., 1983). Considering this, a crucial task for cognitive science is to identify those conditions under which people can resist the influence of heuristics and instead draw upon logical principles to guide their reasoning.

The belief bias effect in reasoning is a foundational effect that has underpinned the development of traditional dual process accounts of reasoning (Evans, 2003; Evans & Stanovich, 2013; Kahneman, 2011; Stanovich & West, 2000). According to these accounts, belief-based responses are generated quickly, drawing upon intuitive 'Type 1' processes, and when in conflict with a logical response, can lead to error unless reasoners have the capacity and motivation to apply logical rules through deliberative 'Type 2' processing (Evans & Curtis-Holmes, 2005). Belief bias is assumed to be the result of a human miserliness, or the inclination to restrict cognitive effort, rather than to engage in effortful logical reasoning (Stanovich, 2018; Toplak et al., 2014).

Recently, researchers have challenged this view, claiming that reasoning based upon logical principles can, in fact, occur at a fast and intuitive level of processing. Dual process 2.0 models (De Neys, 2018; De Neys & Pennycook, 2019) claim that belief-based and logic-based responses are activated intuitively and in parallel, leading to the detection of conflict that may or may not be resolved through the intervention of Type 2 deliberative processes (e.g., Bago & De Neys, 2017; De Neys, 2012; Handley & Trippas, 2015). This claim is supported by evidence that reasoning under conflict tends to take longer, causes a decrease in confidence ratings (De Neys & Glumicic, 2008; De Neys et al., 2011) and feelings of rightness (Thompson et al., 2011), and increases in feelings of error (Gangemi et al., 2015). These effects hold even when reasoners give a biased belief-based response, suggesting that the logical response is activated, irrespective of whether this response is given.

The evidence for 'logical intuition', is not confined to studies in which participants are instructed to provide a logical response. In recent research, reasoners have been asked to judge how much they 'like' a presented statement that might or might not logically follow from the statements presented immediately preceding it (Morsanyi & Handley, 2012; Trippas et al., 2016). In other studies, participants may be asked to rate the brightness (contrast) of the conclusion (Ghasemi et al., 2022[a]), or to generate random responses to a series of logical problems (Howarth et al., 2022). Logical validity has been shown to systematically influence liking and brightness ratings and influence random responses.

Despite the diverse evidence supporting the 'logical intuition' hypothesis, researchers have recently challenged the claim that the effects of logic arise from automatic intuitive processing. For example, research has shown that those higher in cognitive ability produce stronger logic effects in their liking judgments (Hayes et al., 2020; Ghasemi et al., 2022[a]), and that the logic effect is moderated by working memory load (Hayes et al., 2020). More recently, Meyer-Grant and colleagues showed that the liking logic effect is stronger for participants who self-report utilising logic in their judgments (Meyer-Grant et al., 2022). This suggests that logic effects on these tasks arise because the requirements of the task are unclear and participants are hence utilising logical structure as a cue for responding. Similarly, with regards to the brightness judgment task, logic effects only emerge when the brightness discrimination is difficult and it has been argued that under such challenging judgment conditions, participants utilise the logic cue as a basis for responding (Hayes et al., 2022).

The claim that intuitive logic arises from deliberative processes is so far limited to the liking and brightness judgment tasks. However, it is possible that the conflict detection effects described earlier do not arise because the logical inference is activated automatically, but instead because reasoners are engaged in deliberative reasoning. After all, a typical conflict reasoning paradigm involves explicitly instructing participants to reason logically and then evaluating the impact of belief-logic conflict on a range of behavioural measures such as accuracy, latency, and confidence. In this study we utilise an instructional manipulation in which participants are instead presented with conflict problems and asked to either judge whether the conclusion is believable or whether it is logically valid. Previous studies utilising this task have shown that belief-logic conflict interferes with both belief judgments and logic judgments (Handley et al., 2011; Howarth et al., 2016; 2019; Trippas et al., 2017). This suggests that under the belief instruction condition, reasoners are automatically drawing the logical inference and this is then interfering with their ability to evaluate the belief status of the conclusion. The current study uses pupil dilation and eye gaze measures to determine whether conflict is detected and, hence go some way to determining whether intuitive logic effects reflect intuitive processing. Before outlining our study in more detail, we briefly review relevant literature on these eye movement measures.

The study of cognitive processes and pupillometry, dating as far back as the 1960s (Hess & Polt, 1964; Kahneman & Beatty, 1966), has shown that an increase in pupil diameter occurs in response to an increase in the demands of a task. Van der Wel and van Steenbergen's (2018) review of cognitive control tasks (updating, task switching, and inhibition) confirmed that pupil dilation can be used as an indirect index of [cognitive] effort. For example, pupil dilation is related to inhibitory control (Hershman & Henik, 2018), conflict monitoring (van Steenbergen & Band, 2013), and cognitive control (Cavanagh et al., 2014). Crucially, recent research has also revealed that pupil-linked arousal is positively related to uncertainty in serial choice tasks (Urai et al., 2017) and negatively correlated with metacognitive confidence in a decision task (Lempert et al., 2015). Amid this resurgence of pupillometry techniques, reasoning research demonstrated that both 1) metacognitions like confidence and feelings of rightness (for a review see, Ackerman & Thompson, 2017) and 2) judgments of brightness varied as a function of belief-logic conflict (Trippas et al., 2016; Ghasemi et al., 2022[a]). Together these findings provide a strong rationale for expecting pupil dilation to vary as a function of belief-logic conflict.

In addition to pupil dilation, eye movements may also be sensitive to belief-logic conflict. Previous research has demonstrated that some forms of conflict during reasoning affects reasoner's eye movements; Ball (2014) found that participants took longer examining conflict problems even when they give belief based responses in error (Ball et al., 2006). Purcell and colleagues (2022) demonstrated that conflict versions garnered a greater number of eye movements than no-conflict versions of the Cognitive Reflection Test (CRT) supporting the hypothesis that logical responding was the result of intuitive rather than reflective cognitive pathways. Moreover, recent studies reveal that eye movements are negatively related to judgments of confidence (Purcell et al., 2022; in press) and positively related to the engagement of effortful thinking (Purcell et al., in press). The growing evidence suggesting that metacognitions and conflict have systematic impacts on eye movements is promising for the use of gaze-based indicators in determining the intuitive impacts of belief-logic conflict.

Given the relationships between conflict processing and both pupil dilation and gaze, we expected these physiological markers to be sensitive to the detection of conflict under logical instructions, where an effect of conflict arises from competing beliefs. Importantly, however, this study also addresses the more controversial claim that we should also observe physiological indications of conflict under belief instructions, where an effect of conflict arises because of the activation of a competing logical response. In summary, across both logic and belief instructions, we expected to find differences in these markers for problems in which logical structure and believability are aligned (no conflict) or misaligned (in conflict). Specifically, we predicted that:

Hyp. 1: Lower accuracy, increased response latency, increased pupil dilation, and a greater number of gaze movements would be observed on conflict compared to no conflict problems across both logical and belief instructions.

In addition to our primary comparison of conflict and no conflict problems, we also explored whether reasoners showed sensitivity to conflicting logic when they generated belief-based responses (either in error under logic instructions, or correctly under belief instructions). If participants give a belief-based response on conflict problems under logical instructions in error, but the logical response is activated, we would expect this to be marked by increased latency, gaze shifts and dilation compared to non-conflict problems. By a similar rationale, if participants correctly give a correct belief-based response under belief instructions on conflict problems, if the competing logical response is activated, then they should show increased latency, gaze and dilation compared to trials in which the logical and belief-based responses are aligned. Following this argument, we also predicted:

Hyp. 2: Under belief instructions, participants will be slower, show a greater number of eye movements, and larger pupil dilation on correct conflict trials compared to correct no conflict trials.

Hyp. 3: Under logic instructions, participants will be slower, show a greater number of eye movements, and larger pupil dilation on incorrect conflict trials (i.e., belief-aligned responses with logically conflicting information) compared to correct no conflict trials.

2. Method

2.1. Participants

Thirty-eight undergraduate psychology students participated in the experiment in exchange for course credit. Ages ranged from 18 to 36 (*M*=19.76, *SD*=3.54), with 27 participants identifying as female and 11 identifying as male. All participants had normal vision (i.e., no glasses or contact lenses).

2.2. Procedure

The chair and chinrest were adjusted to comfortable positions for the participant before commencing the experiment. After a nine-point eye-tracking calibration, participants were presented with instructions (see Supplementary Material) and four practice trials. Experimental trials were presented in four blocks with a three-minute break between each block. A nine-point calibration was conducted before each block of trials and a one-point calibration was conducted before each item. If a one-point calibration failed, a nine-point calibration was repeated.

At the start of each trial, a fixation cross appeared in the centre of the screen for 3000ms. After this, the major premise appeared for 3000ms, followed by the minor premise for 3000ms, and then the conclusion. After the conclusion was displayed for 3000ms, it was joined by the two response options (valid/invalid or believable/unbelievable). The conclusion and response options remained on screen until the participant made their response. This means that participants only received the instruction on how to respond *after* being presented with the reasoning problem. At the end of the experiment, participants provided their age and gender.

2.3. Apparatus and Materials

Reasoning task. Participants were presented with 96 reasoning problems in a random order. Half the problems were conditional arguments and half were disjunctions. For each problem type, half were logically valid, and half were determinately invalid. For each level of validity, half the arguments had believable conclusions and half had unbelievable conclusions. Validity and believability were then crossed to determine the conflict-status of the argument, such that valid and believable or invalid and unbelievable resulted in a non-conflict trial, whereas valid and unbelievable or invalid and believable resulted in a conflict trial. Finally, for each level of conflict, half the arguments were paired with the instruction to answer according to logical validity, and half were paired with the instruction to answer according to conclusion believability. The design was a 2 (belief versus logic instructions) by 2 (conflict vs no-conflict) by 2 (conditional vs disjunctive problems) fully within-participants (See the supplementary materials for the full set of arguments). Accuracy and response latencies were recorded for every trial.

Eye and pupil tracking equipment. The reasoning task was built using Experiment Builder presentation software 1.10.165 (SR Research) and presented on a 24.5-inch LCD monitor (240 Hz, 1920 x 1080). Eye movements and pupil dilation were recorded monocularly (right eye) with a desk mounted EyeLink 1000 tracker sampling at a rate of 1000 Hz (SR Research). A chinrest was used to stabilise head movements and maintain viewing distance of 800mm. Responses were made by pressing the left (valid/believable) or right (invalid/unbelievable) key on a Cedrus box.

Eye and pupil tracking measures. Three Areas Of Interest (AOIs) were created for the conclusion and response screen (see Figure 1). As shown in Figure 1, AOIs were created for 1) the conclusion presented in the centre of the screen, 2) for the 'left' option presented below the conclusion, equidistant from the centre and the left edge of the screen, and 3) for the 'right' option presented below the conclusion, equidistant from the centre and the right edge of the screen. For our analysis, saccades were calculated as the number of eye movements that occurred between any two of the three AOIs whilst the conclusion and response options remained on screen (Purcell et al., 2022). A pupil measurement was recorded every 20ms provided the participant was fixating within AOI 1 and the eye was not in saccade[1]. For our analysis, pupil dilation was calculated as the average dilation measurement from when the response options were presented until either a response was recorded or until 2000ms after the response options were presented.

[1] To prevent extreme viewing angles distorting the pupil measurements, only pupil measurements that occurred when the participant was fixating, that is, the eye was not in movement, and when they were looking within the 'conclusion' AOI were included in the analysis.

Figure 1. Participants were presented with each item gradually such that they saw a fixation cross, premise 1, premise 2, the conclusion, and finally, the conclusion and response options. This figure represents the final stage of the item presentation with the conclusion and response items on screen. Saccades were calculated as any eye movements between two of the AOIs and pupil measures were taken while the participant was fixating within AOI 1.

3. Results

3.1. Analysis Plan

The following analyses are divided into three sections. In section one, we report our primary analyses concerning the effects of conflict, instructions, and problem-type, on accuracy, latency, gaze, and dilation. In sections two and three, we report our exploratory analyses focusing on belief instructions and logic instructions, respectively. All analyses for latency, gaze, and dilation used general linear mixed models[2], whereas the analysis for accuracy used a generalised mixed model with a logit function[3]. Raw data and scripts for the analysis analyses are stored on OSF:

[2] All linear models estimated using REML and nloptwrap optimizer.

[3] All analyses were conducted in RStudio (RStudio Team, 2022) using the lme4 package (Bates et al., 2015). Models were estimated using ML and Nelder-Mead optimizer.

https://osf.io/tsbp5/?view_only=fe153d8d4b74480587584c7484b3d2ea

3.1.1. Section 1

Here, we assess the effects of three key independent variables – conflict (conflict, no conflict items), instructions (belief, logic), and problem type (conditional, disjunctive) – on our four key dependent variables: accuracy, latency, gaze patterns, and pupil dilation. See Table 1 for model outputs and Table A1 for extended descriptive results.

The model predicting accuracy showed a significant three-way interaction. As expected, accuracy was lower for conflict (*M*=31.70, *SE*=1.15) than no conflict items (*M*=44.25, *SD*=.45) for all combinations of instructions and problem type. However, this effect was strongest for disjunctives under belief instructions, followed by conditionals under belief instructions, then conditionals under logic instructions, and finally disjunctives under logic instructions.

The model predicting latency showed main effects of conflict and instructions. Latencies were longer for conflict (*M*=1161.59, *SE*=13.16) than no conflict items (*M*=1004.17, *SE*=12.51), and longer for belief instructions (*M*=1148.18, *SE*=13.42) than logic instructions (*M*=1017.58, *SE*=12.31). No other effects were significant.

The model predicting gaze showed significant main effects of conflict and instructions. More gaze movements occurred between the three AOIs (conclusion and left and right options) for conflict (*M*=1.59, *SE*=0.04) than no conflict items (*M*=1.20, *SE*=.03), and for belief (*M*=1.56, *SE*=.04) than logic instructions (*M*=1.22, *SE*=.03). No other effects were significant.

The model predicting dilation revealed a main effect of conflict together with a significant three-way interaction between conflict, instructions, and problem type. Pupil dilation was significantly greater for conflict than no conflict items for conditionals under belief instructions and for disjunctives under validity instructions and marginally significant for conditionals under logic instructions (means and variances relevant for three-way interactions are presented in Table A1). The effect of conflict for disjunctives under belief instructions was not significant, but the means trended in the same direction.

In relation to Hypothesis 1, these findings show a largely consistent effect of conflict in the predicted direction; participants were less accurate, slower, showed more saccades, and larger pupil dilation on conflict compared to no conflict items.

Table 1
The effects of conflict, instructions, and problem type, on accuracy, latency, gaze, and dilation

	Accuracy		Latency		Gaze		Dilation	
Effect [Level]	*b*	*CI*	*b*	*CI*	*b*	*CI*	*b*	*CI*
Intercept	**2.40****	**2.04, 2.75**	-.03	-.20, .13	-.04	-.18, .10	.02	-.28, .31
Conflict [Conflict]	**-1.87****	**-2.23, -1.52**	**.29****	**.19, .40**	**.26****	**.14, .37**	**.08****	**.03, .12**
Instructions [Validity]	.20	-.24, .65	**-.19****	**-.30, -.09**	**-.13***	**-.24, -.01**	-3.20e-03	-.05, .04
Problem Type [Logic]	.45	-.02, .92	-.02	-.13, .10	2.26e-03	-.11, .12	-.02	-.06, .03
Instructions x Conflict	.08	-.43, .60	-.06	-.21, .09	-.15	-.31, .01	-.04	-.10, .02
Problem Type x Conflict	**-.81***	**-1.35, -.28**	.05	-.10, .20	.02	-.14, .19	**-.07***	**-.13, -6.97e-03**
Problem Type x Instructions	-.53	-1.18, .12	.07	-.08, .22	.01	-.15, .17	-.04	-.10, .02
Problem Type x Conflict x Instructions	**1.59****	**.83, 2.35**	-.11	-.32, .10	.07	-.16, .30	**.09***	**6.19e-03, 0.18**

Note: Significant effects are in bold. **p*<.05, ***p*<.001.

3.1.2. Section 2

Here we present our secondary analyses examining items presented with 'belief' instructions and assess whether the indicators – latency, gaze, and dilation – are impacted by trial type (specifically, correct conflict compared to correct no conflict trials) and problem type (conditional, disjunctive). By comparing our observations between a) correct responses made when logical structure is consistent with the belief response (trial type: correct no conflict), and b) correct belief responses made when the logical structure is inconsistent with the belief response (trial type: correct conflict), we can assess whether the conflicting logical information impacts accurate belief-based reasoning. See Table 2 for model output and Table B1 for extended descriptive results.

The model predicting latency showed a significant two-way interaction between trial and problem type. Latencies were greater for correct conflict than correct no conflict items; however, this effect was stronger for disjunctive arguments instructions (means and variances relevant for two-way interactions are presented in Table B1).

The model predicting gaze showed significant main effects of trial and problem type. Saccades between AOIs (question, two response options) were greater for correct conflict (*M*=2.06, *SE*=.08) than correct no conflict (*M*=1.22, *SE*=.04), and for disjunctives (*M*=1.65, *SE*=.07) than conditionals (*M*=1.44, *SE*=.06).

The model predicting dilation showed a significant two-way interaction between trial and problem type. Dilation was greater for correct conflict than correct no conflict trials, for conditionals but no difference was observed for disjunctives (see Table B1).

In line with Hypotheses 2, Section 2 observed consistent main effects of trial type, showing that participants were less accurate, slower, showed more saccades, and larger pupil dilation on correct conflict trials compared to correct no conflict trials. The findings provide evidence of logical structure impacting belief-based responding. Whilst the effect of conflict on disjunctive arguments was significant for latency and gaze, it was not significant for dilation.

Table 2
The effects of trial type and problem type on latency, gaze, and dilation under belief instruction.

	Latency		Gaze		Dilation	
Effect	*b*	*CI*	*b*	*CI*	*b*	*CI*
Intercept	-0.03	-0.19, 0.12	-0.03	-0.16, 0.09	0.04	-0.24, 0.31
Trial Type	**-0.18****	**-0.24, -0.12**	**-0.16****	**-0.22, -0.09**	**-0.04***	**-0.06, -0.01**
Problem Type	**0.07***	**-0.01, 0.16**	**0.12***	**0.03, 0.21**	**-0.03***	**-0.06, 3.88e-03**
Trial Type* Problem Type	**-0.09***	**-0.18, -5.78e-03**	-0.08	-0.17, 0.01	**0.03***	**1.58e-03, 0.07**

Note: Significant effects are in bold. *p<.05, **p<.001.

3.1.3. Section 3

In Section 3, we present our secondary analyses, examining items presented under 'logic' instruction, and assess whether latency, gaze, and dilation are impacted by trial type (specifically, incorrect conflict compared to correct no conflict) and problem type (conditional, disjunctive). That is, we compare belief-based responses in cases where there is conflicting logical structure (trial type: incorrect conflict) to belief-based responses with consistent logical structure (trial type: correct conflict). This allows us to evaluate whether there is evidence that reasoners detect conflict when they give an erroneous belief-based response. See Table 3 for model output and Table C1 for extended descriptive results.

The model predicting latency showed a significant effect of trial type; latencies were longer for incorrect conflict trials (*M*=1146.36, *SE*=37.19) than correct no conflict trials (*M*=915.61, *SE*=16.74).

The model predicting gaze showed a significant effect of trial type; a greater number of saccades were observed on incorrect conflict trials (*M*=1.34, *SE*=.10) than correct no conflict trials (*M*=1.00, *SE*= .04).

The model predicting dilation showed a marginal effect of trial type in the expected direction; dilation was greater for incorrect conflict trials (*M*=778.03, *SE*=18.28) than correct no conflict trials (*M*=721.48, *SE*=7.76).

The findings show support for Hypothesis 3. We observed that participants took longer and showed more saccades on incorrect conflict compared to correct no conflict trials. The marginal effect of trial type on dilation was in the expected direction, with greater dilation for incorrect conflict trials.

Table 3
The effects of trial type and problem type on latency, gaze, and dilation under logic instructions

	Latency		Gaze		Dilation	
Effect	*b*	*CI*	*b*	*CI*	*b*	*CI*
Intercept	0.11	-0.05, 0.27	0.03	-0.11, 0.17	0.04	-0.24, 0.33
Trial Type	**-0.17****	**-0.24, -0.10**	**-0.08***	**-0.15, -7.10e-03**	**-0.02±**	**-0.05, 4.25e-03**
Problem Type	1.05	-0.04, 0.17	0.06	-0.04, 0.17	-0.03	-0.07, 5.27e-03
Trial Type* Problem Type	-4.80e-03	-0.11, 0.10	-0.06	-0.16, 0.05	-0.03	-0.07, 0.01

Note: Significant effects are in bold. * p<.1, *p<.05, **p<.001.

4. Discussion

A core claim of Dual Process 2.0 models is that a reasoner can generate intuitions based upon both beliefs and logical rules; therein, these models challenge the long-held view that logical reasoning is the exclusive domain of effortful, Type 2 processes. Further, they claim that competition between any two intuitive responses (including those stemming from logical structure) underpins the mechanism through which Type 2 processes are engaged and conflict can be resolved (DeNeys, in press). Using a comprehensive range of behavioural and physiological measures, we provide evidence in support of the Dual Process 2.0 challenge and, specifically, for the hypothesis that reasoners are intuitively sensitive to logical structure and can automatically activate a logical response when engaged in reasoning under conflict.

In this study we evaluated the intuitive nature of logical intuitions by examining the impact of conflict on measures of gaze and pupil dilation alongside behavioural measures of accuracy and latency. We also employed the instructional manipulation paradigm allowing us to examine the impact of conclusion validity on belief judgments, an effect that has been argued to arise because the logical inference is automatically activated and interferes with the generation of a belief-based response. The findings of our primary analyses were relatively clear cut; We replicated earlier research showing that conflict impacts accuracy and latency under both logical and belief-based instructions and overall belief judgments take longer than logic judgments. The effect of conflict under belief instructions and the longer latency for these judgments provides support for the claim that a competing logical inference is being cued. In addition, the effect of conflict extended to our physiological measures resulting in an increased number of gaze shifts (between the conclusion and two response options) together with increased pupil dilation.

Typically, research on conflict detection evaluates the impact of conflict on confidence and latency when participants are instructed to respond logically but instead respond based upon beliefs. Lower confidence and longer latency are indications that participants have activated the logical inference, despite giving a biased response. The use of belief instructions in this study allowed us to examine whether participants who correctly give a belief-based response also show evidence of intuitively activating the logical response, despite responding correctly on a conflict trial. Our secondary analyses showed, consistent with earlier work, that incorrect conflict trials under logical instructions are associated with longer latencies, more gaze shifts and marginally greater pupil dilation than matched non-conflict trials. Importantly these findings extend to the comparison of correct conflict and non-conflict trials under belief instructions, where we observe longer latencies and more gaze shifts respectively. Pupil dilation also differs on conditional arguments, but this effect did not extend to disjunctions. Taken together these findings provide compelling collective evidence that reasoners are activating a competing logical response irrespective of whether they give an incorrect (under logical instructions) or correct (under belief instructions) belief-based response under conflict.

The broad convergence of findings across our behavioural and physiological measures is consistent with the claim that the detection of structural problem features that align with logical validity is automatic in nature. That is, reasoners activate the competing response rapidly and without awareness. This claim contrasts with recent research which shows that the effects of logic on liking tasks is influenced by working memory constraints, linked to cognitive capacity, and correlated with explicit logic effects (Ghasemi et al., 2022[a]; Hayes et al., 2020). We agree with arguments that the liking task may not be an ideal measure of intuitive logic, as the instructions lack clarity as to the basis on which reasoners should make their judgments. Hence, it is unsurprising that some participants might explicitly draw upon logical features to inform their judgment. The advantage of the manipulation employed in this study is that participants are explicitly instructed to respond based upon beliefs, and as our analyses clearly show, there remains very clear evidence that they activate a competing logical response even when responding successfully.

Does this mean that the logical response is activated intuitively? The combined evidence from this study suggests that this is, in fact, the case. Participants show longer latencies, an increased number of gaze shifts between the conclusion and response options, and some evidence of greater pupil dilation on a subset of the problems. The fact that reasoners who provide an accurate response on conflict items nevertheless show physiological markers of conflict detection is consistent with the automatic activation of a competing response. Research suggests that pupil dilation may reflect engagement of mental effort, particularly that associated with conflict detection, and more recently it has been associated with metacognitive processes liked to uncertainty and reduced confidence (Lempert et al., 2015). This is consistent with research which shows that the presence of belief-logic conflict leads to reduced confidence and lower feelings of rightness (Thompson et al., 2011). Recent research shows that perceptual ambiguity can lead to pupil dilation even when the observer remains consciously unaware of this ambiguity (Graves et al., 2021). A task for future research will be to examine whether reasoners are similarly unaware of a competing logical response under conflict when they give belief-based responses on this task.

Finally, we note recent research which claims that logical intuitions of the kind that we describe here do not arise because of the activation of logical rules, but instead because of more superficial structural cues in the problem (Meyer-Grant et al., 2022). These authors show that logic effects on liking judgment tasks arise because on valid arguments there is a match between the polarity of the premises and the conclusions, which is not present on invalid versions of the same arguments. Their explanation draws upon an ‘atmosphere heuristic’ and not logic as the principal cue that that determines liking evaluations. This claim aligns with a similar ‘matching’ account developed in recent research by Ghasemi et al. (2022[b]) who argued that the effect of validity on belief judgments similarly arose because of a polarity match on valid arguments, an effect that was extended to invalid conditional arguments where the same structural features were present. Interestingly, in both studies, the same non-logical features influenced explicit logical judgments, a finding that aligns with classic work in the field of deductive reasoning (Evans et al., 1993; Wetherick & Gilhooly, 1995).

Whilst our current work cannot directly discriminate between these heuristic-based accounts of logical intuition effects, we note that our study also included disjunctive arguments, where a matching heuristic of the kind described above cannot apply (the polarity of the conclusion and the premises mismatch). Although there were some minor differences in the effects observed between conditionals and disjunctives, the patterns of conflict detection across the majority of measures were similar for both connectives. It is, of course, possible that a different heuristic operates with disjunctive arguments, and we recognise that future research might establish that logical intuitions of the kind we have examined in this paper do, in fact, also arise through non-logical processes. Irrespective of this, the current research demonstrates clearly that competing intuitions (whether based on logic or heuristics) are activated rapidly whilst reasoning and their co-occurrence is evidenced by both behavioural and physiological markers.

5. References

Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21*(8), 607–617.

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

Ball, L. J. (2014). Eye-tracking and reasoning: What your eyes tell about your inferences. In W. De Neys & M. Osman (Eds.), *New approaches in reasoning research* (pp. 51–69). Psychology Press.

Ball, L. J., Lucas, E. J., Miles, J. N. V, & Gale, A. G. (2003). Inspection Times and the Selection Task: What do Eye-Movements Reveal about Relevance Effects? *Quarterly Journal of Experimental Psychology A. Human Experimental Psychology.*, *56*(6), 1053–1077.

Ball, L. J., Phillips, P., Wade, C. N., & Quayle, J. D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing models. *Experimental Psychology, 53*(1), 77–86.

Bates D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General, 143*(4), 1476–1488.

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

De Neys, W. (2018). Dual process theory 2.0. Routledge.

De Neys (in press). Advanced theorising about fast and slow thinking. *Brain and Behavioural Sciences.*

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*(3), 1248–1299.

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science, 28*(5), 503–509.

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: conflict and decision confidence. *PLoS One, 6*(1), e15954.

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Evans J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology*, 59, 255–278.

Evans, J. St. 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.

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

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

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Lawrence Erlbaum Associates, Inc.

Gangemi, A., Bourgeois-Gironde, S., & Mancini, F. (2015). Feelings of error in reasoning—in search of a phenomenon. *Thinking & Reasoning, 21*(4), 383–396.

Ghasemi, O., Handley, S., & Howarth, S. (2022)[a]. The bright homunculus in our head: Individual differences in intuitive sensitivity to logical validity. *Quarterly Journal of Experimental Psychology, 75*(3), 508-535.

Ghasemi, O., Handley, S.,Howarth, S., Newman, I. R., & Thompson, V. A. (2022)[b]. Logical Intuition is not really about logic. *Journal of Experimental Psychology: General, 151,* 2009-2028.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge University Press.

Graves, J. E., Egre, P., Pressnitzer, D., & de Gardelle, V. (2021). An implicit recognition of stimulus ambiguity in pupil size. *Proceedings of the National Academy of Sciences, 118.* E210799718.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27

Handley, S., & Trippas, D. (2015). Dual processes, knowledge, and structure: A critical evaluation of the default interventionist account of biases in reasoning and judgment. *Psychology of Learning and Motivation, 62*, 33-56.10.

Handley, S. J., Newstead, S. E., & Trippas, D. (2011). Logic, beliefs, and instruction: a test of the default interventionist account of belief bias. *J Exp Psychol Learn Mem Cogn, 37*(1), 28-43.

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.

Hayes, B. K., Stephens, R. G., Lee, M. D., Dunn, J. C., Kaluve, A., Choi-Christou, J., & Cruz, N. (2022). Always look on the bright side of logic? Testing explanations of intuitive sensitivity to logic in perceptual tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. *Behavior Research Methods, 50*(1), 107–114.

Hess, E.H. and Polt, J.M. (1964) Pupil Size in Relation to Mental Activity during Simple Problem Solving. *Science, 140*, 1190-1192.

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.

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.

Howarth, S., Handley, S., & Polito, V. (2022). Uncontrolled logic: intuitive sensitivity to logical structure in random responding. *Thinking & Reasoning, 28*(1), 61-96.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Kahneman, D., & Frederick, S. (2005). A Model of Heuristic Judgment. In K. Holyoak & R. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 267–293). Cambridge University Press.

Lempert KM, Chen YL, Fleming SM (2015) Relating Pupil Dilation and Metacognitive Confidence during Auditory Decision-Making. PLoS ONE 10(5): e0126588.

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin, 126*(1), 109–137.

Meyer-Grant, C. G., Cruz, N., Singmann, H., Winiger, S., Goswami, S., Hayes, B. K., & Klauer, K. C. (2022). Are logical intuitions only make-believe? Reexamining the logic-liking effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences, 14*(10), 435–440.

Morsanyi, K., & Handley, S. J. (2012). Logic feels so good-I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *J Exp Psychol Learn Mem Cogn, 38*(3).

Purcell, Z. A., Wastell, C. A., Howarth, S., Roberts, A. & Sweller, N. (2022). Eye tracking and the cognitive reflection test: Evidence for intuitive correct responding and uncertain heuristic responding. *Memory & Cognition, 50(2),* 348-365.

Purcell, Z., Wastell, C. A., & Sweller, N. (in press): Eye Movements Reveal that Low Confidence Precedes Deliberation. *Quarterly Journal of Experimental Psychology*.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219–235.

RStudio Team. (2022). *RStudio: Integrated Development for R*. Rstudio, PBC, Boston, MA. http://www.rstudio.com/.

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

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63*(3), 107–140.

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.

Trippas, D., Handley, S. J., Verde, M. F., & Morsanyi, K. (2016). Logic brightens my day: Evidence for implicit sensitivity to logical validity. *J Exp Psychol Learn Mem Cogn, 42*(9), 1448-1457.

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.

Urai AE, Braun A & Donner TH. (2017) Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, *8.* 14637.

van Steenbergen, H., & Band, G. P. H. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience, 7,* Article 215

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review, 25*(6), 2005–2015.

Wetherick, N., & Gilhooly, K. (1995). 'Atmosphere', matching, and logic in syllogistic reasoning. *Current Psychology, 14*(3), 169-178.

6. Appendices

6.1 Appendix A

A1. In Section 1.1, we observed that accuracy was predicted by a three-way interaction effect of conflict (conflict, no conflict), instructions (belief, logic), and problem type (conditional, disjunctive) on accuracy; follow up analyses showed that:

- Under believability instructions for conditionals, the effect of conflict is statistically significant and negative (*b* = -2.05, 95% CI [-2.43, -1.67], $p < .001$)
- Under believability instructions for disjunctives, the effect of conflict is statistically significant and negative (*b* = -2.99, 95% CI [-3.44, -2.54], $p < .001$)
- Under validity instructions for conditionals, the effect of conflict is statistically significant and negative (*b*= -2.03, 95% CI [-2.44, -1.61], $p < .001$)
- Under validity instructions for disjunctives, the effect of conflict is statistically significant and negative (*b*= -1.11, 95% CI [-1.51, -0.71], $p < .001$)

Table A1

This table provides the means and standard errors for each combination of our primary variables: conflict, instructions, and problem type. In line with our expectations, we observe lower means for conflict items on accuracy, but higher means on conflict items for latency (ms), gaze (saccades), and pupil dilation.

Conflict	Instructions	Problem Type	Accuracy M (SE)	Latency M (SE)	Gaze M (SE)	Dilation M (SE)
No Conflict	Believability	Conditional	10.89 (0.17)	1054.34 (26.38)	1.30 (0.07)	733.68 (11.17)
		Disjunction	11.27 (0.14)	1049.07 (25.84)	1.33 (0.07)	729.38 (10.87)

	Validity	Conditional	11.09 (0.14)	939.61 (23.12)	1.08 (0.06)	733.64 (10.84)
		Disjunction	11.00 (0.21)	973.68 (24.34)	1.10 (0.06)	720.60 (10.73)
Conflict	Believability	Conditional	7.52 (0.55)	1231.67 (27.01)	1.80 (0.09)	752.66 (11.4)
		Disjunction	6.52 (0.56)	1257.67 (26.87)	1.82 (0.08)	732.59 (10.8)
	Validity	Conditional	8.07 (0.55)	1078.53 (24.98)	1.27 (0.07)	742.82 (11.05)
		Disjunction	9.59 (0.38)	1078.49 (25.43)	1.46 (0.11)	735.28 (11.34)

6.2. Appendix B

6.2.1. B1. In Section 2.1, we observed that – for items presented under belief instructions – latency was predicted by a two-way interaction of trial type (correct no conflict, correct conflict) and problem type (conditional, disjunctive; see Table B1). Follow up analyses showed that:

- For conditionals, the effect of trial type is statistically significant and negative (t(763) = -6.00, *p* < .001, *b* = -0.19, 95% CI [-0.26, -0.13])
- For disjunctives, the effect of trial type is statistically significant and negative (t(738) = -8.62, *p* < .001; *b*= -0.27, 95% CI [-0.34, -0.21])

Table B1

This table provides the means and standard errors for each combination of our trial type (correct conflict, correct no conflict) and problem type (conditional, disjunctive) under belief instructions. We observe higher means on conflict items for latency (ms) and gaze (saccades) for both conditionals and disjunctives, however, for dilation, this pattern only holds for conditionals.

Trial Type	Problem Type	Latency M (SE)	Gaze M (SE)	Dilation M (SE)
Correct No Conflict	Conditional	1010.58(26.86)	1.17(0.06)	732.61(11.64)
	Disjunction	1017.11(26.1)	1.28(0.07)	731.25(11.33)
Correct Conflict	Conditional	1306.36(33)	1.84(0.12)	760.02(16.55)
	Disjunction	1450.6(31.86)	2.32(0.12)	724.67(16.79)

6.2.2. B2. In section 2.3, we observed that – for items presented under belief instructions – dilation was predicted by a two-way interaction between trial type (correct no conflict, correct conflict) and problem type (conditional, disjunctive). Follow up analyses showed that:

- For conditionals, the effect of trial type is statistically significant and negative (t(761) = -2.88, *p* =.004; *b* = -0.03, 95% CI [-0.06, -0.01]).
- For disjunctives, the effect of trial type is statistically non-significant and negative (t(736) = -0.70, *p* = 0.483; *b* = -8.94e-03, 95% CI [-0.03, 0.02]).

6.3. Appendix C

Table C1

This table provides the means and standard errors for each combination of trial type (incorrect conflict, correct no conflict) and problem type (conditional, disjunctive) under logic instructions. We observe higher means on conflict items for latency (ms), gaze (saccades) and dilation, however, the difference is not significant for dilation.

Trial Type	Problem Type	Latency M (SE)	Gaze M (SE)	Dilation M (SE)
Correct No Conflict	Conditional	898.61 (22.85)	0.98 (0.06)	726.99 (10.89)
	Disjunction	932.75 (24.48)	1.02 (0.06)	715.93 (11.07)
Incorrect Conflict	Conditional	1124.17 (47.58)	1.24 (0.12)	788.20 (23.68)
	Disjunction	1182.18 (59.73)	1.51 (0.19)	761.71 (28.76)