Psychophysiology, 51 (2014), 290–297. Wiley Periodicals, Inc. Printed in the USA.
Copyright © 2013 Society for Psychophysiological Research
DOI: 10.1111/psyp.12169

Heuristic and analytic processes in reasoning: An event-related potential study of belief bias

ADRIAN P. BANKS AND CHRISTOPHER HOPE
School of Psychology, University of Surrey, Guildford, UK

Abstract

Human reasoning involves both heuristic and analytic processes. This study of belief bias in relational reasoning investigated whether the two processes occur serially or in parallel. Participants evaluated the validity of problems in which the conclusions were either logically valid or invalid and either believable or unbelievable. Problems in which the conclusions presented a conflict between the logically valid response and the believable response elicited a more positive P3 than problems in which there was no conflict. This shows that P3 is influenced by the interaction of belief and logic rather than either of these factors on its own. These findings indicate that belief and logic influence reasoning at the same time, supporting models in which belief-based and logical evaluations occur in parallel but not theories in which belief-based heuristic evaluations precede logical analysis.

Descriptors: Dual process theory, Belief bias, ERP, P3, N2

The idea that there are two kinds of thinking has a long history within psychology (e.g., James, 1890/1950). One popular distinction is between "intuitive" and "analytic" thought. Theoretical accounts advancing this idea are referred to collectively as dual process theories (Evans, 2008). Recently, dual process theories have been successful in explaining a range of empirical phenomena in reasoning and decision making (e.g., Evans & Over, 1996; Kahneman & Frederick, 2002; Stanovich & West, 2000). Dual process theories all rely on two types of processing that vary in a number of properties, most importantly whether they make high or low demands on working memory (Evans, 2008). These will be referred to here as heuristic and analytic processes (e.g., Evans, 2006). A number of methods have been used to demonstrate the presence of two types of reasoning process (e.g., De Neys, Moyens, & Vansteenwegen, 2010; De Neys, Vartanian, & Goel, 2008). However, a key theoretical question remains unresolved: How do these processes interact?

Several models have been suggested (Evans, 2007). For the purpose of this paper, these will be categorized into two types—serial and parallel—according to the sequence in which the two kinds of thinking occur. Serial models propose that one set of processes occurs prior to the other set. An example is the default-interventionist model (Evans, 2007). In this model, heuristic responses are the default response, but the analytic process may intervene subsequently to alter the response. Parallel models propose that heuristic and analytic processes occur simultaneously. An example is Sloman's dual process model of reasoning (Sloman, 1996). Previous behavioral studies have tested these models, leading to differing conclusions (e.g., Evans & Curtis-Holmes, 2005; Handley, Newstead, & Trippas, 2011). The aim of the present study is to use event-related brain potentials (ERPs) to establish if serial or parallel models provide the most accurate account of dual process theory.

Belief Bias in Reasoning

Belief bias is the tendency, when reasoning, to judge conclusions based on prior beliefs rather than logical validity (e.g., Evans, Barston, & Pollard, 1983; Klauer, Musch, & Naumer, 2000). This bias has been found in a range of reasoning tasks such as syllogistic reasoning (e.g., Evans et al., 1983; Klauer et al., 2000), relational reasoning (e.g., Banks, 2013; Roberts & Sykes, 2003), conditional reasoning (Evans, Handley, & Bacon, 2009), and transitive reasoning (Andrews, 2010). The study of belief bias is a major area of investigation for dual process theories (Evans, 2003).

In a typical belief bias experiment, participants are presented with some premises and asked to evaluate a conclusion; see Table 1 for examples. The logical validity and believability of the conclusion are manipulated independently, resulting in four conditions: believable valid, unbelievable valid, believable invalid, unbelievable invalid. In the first and last condition, there is no conflict between these processes because they both cue the same response. In the middle two conditions, there is a conflict because the two processes cue different responses. Dual process theories of belief bias differ in their explanation of how this conflict is resolved.

Serial models of belief bias predict that a belief-based response is initially cued by a heuristic process, but this can be overridden by a logical response cued by an analytic process (e.g., Stupple, Ball, Evans, & Kamal-Smith, 2011). Parallel models predict that both belief-based and logical processes operate from the start and

Address correspondence to: Adrian P. Banks, School of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK. E-mail: a.banks@surrey.ac.uk

PSYCHOPHYSIOLOGY

Table 1. *Examples of Problems from Each Condition*

Believable valid	
Premise 1	giraffes are bigger than mice
Premise 2	zoots are bigger than mice
Premise 3	elephants are bigger than zoots
Initial conclusion	elephants are bigger than
Final conclusion word	mice
Believable invalid	
Premise 1	elephants are bigger than bicycles
Premise 2	zoots are bigger than bicycles
Premise 3	mice are bigger than zoots
Initial conclusion	elephants are bigger than
Final conclusion word	mice
Unbelievable valid	
Premise 1	giraffes are bigger than elephants
Premise 2	zoots are bigger than elephants
Premise 3	mice are bigger than zoots
Initial conclusion	mice are bigger than
Final conclusion word	elephants
Unbelievable invalid	
Premise 1	elephants are bigger than bicycles
Premise 2	zoots are bigger than bicycles
Premise 3	mice are bigger than zoots
Initial conclusion	mice are bigger than
Final conclusion word	elephants

compete to cue a response. If these responses conflict, then one of them is inhibited (e.g., De Neys, 2012; Handley, Newstead, & Trippas, 2011).

Belief Bias and ERPs

A small number of studies have used ERP techniques to investigate reasoning. The first study of belief bias using ERPs was conducted by Luo et al. (2008). They found an enhanced P500 in problems where belief inhibited the logical response compared to problems where belief facilitated it. However, the experimental design prevents any firm conclusions being drawn about dual reasoning processes. First, the ERPs were time locked to the onset of the minor premise, before the conclusion was presented. Therefore, the believability of the conclusion was irrelevant to the ERPs reported. With this approach, it is not possible to test many current models of belief bias that explain the effect in terms of the influence of the believability of the conclusion. A second concern with this timing is that the ERP measured both reading of the premise (six to eight Chinese characters) and, subsequently, reasoning about it rather than solely reasoning. Third, the study used only invalid problems. Therefore, it is not possible to manipulate logical validity and test the characteristic interaction of logic and belief that is found in studies of belief bias. Finally, as no behavioral effect of belief bias was found, it is not certain that the task elicited the phenomenon.

A subsequent study by Luo et al. (2013) found an increased positive potential 200–400 ms prior to a correct response when belief and logic conflicted compared to when they did not. By time locking to the response rather than the onset of the conclusion, this design does not enable a conclusion to be drawn about whether the initial conclusion processing occurs serially or in parallel. Only their eventual effect on responding can be detected.

An N2 component has also been found on reasoning tasks. The N2 is associated with cognitive control (Folstein & van Petten, 2008). In a study of heuristics in judgment, a differential N2 was found in poor compared to good reasoners, which led the authors to conclude that poor reasoners experienced more difficulty inhibiting biased thinking (De Neys, Novitskiy, Ramautar & Wagemans, 2010).

The Present Study

The present study develops the approach to studying belief bias in order to compare serial and parallel models of dual process theories. We introduce a full factorial design that manipulates both logic and belief, enabling a test of both of these factors. We use a wide range of problems to reduce the likelihood that the format of the problem becomes overlearned and to avoid a ceiling effect masking the behavioral effects of belief bias. Finally, the problems are designed so that it is not possible to evaluate the conclusion until the onset of the final word. ERPs are time locked to this onset. Therefore, the ERPs are more precisely linked to the evaluation of the conclusion, and confounds with other processes are reduced.

We assume that participants read the premises and the first part of the conclusion and form a representation of this information. The critical final word is evaluated within the context of this representation. The differences in ERPs elicited by the final word across conditions indicate the form of the initial representation and the processes influencing its evaluation.

The main predictions of the study concern the timing of the influence of belief and logic. If the serial models of belief bias are correct, the ERPs will initially differ solely as a function of belief because the processes responding to belief are predicted to occur first. If the parallel models are correct, both logic and belief will influence the evaluation at the same time because processes responding to these factors are predicted to occur at the same time. There will initially be an interaction of these two processes, resulting in different ERPs when the two processes generate similar conclusions (no-conflict problems) compared to different conclusions (conflict problems).

Two components are of particular relevance to this task, N2 and P3. N2 amplitude reflects cognitive control and is involved in the resolution of response conflicts (Folstein & van Petten, 2008). The amplitude of N2 is greater when there is a response conflict. In this task, conflict occurs when logic and belief cue different responses. The serial models predict that the initial conclusion processing is purely belief based, and any conflict with logic should occur much later than the typical N2 time window. Therefore, there will be no difference in N2 amplitude between conflict and no-conflict problems if the serial models are correct. Parallel models predict that logic and belief are processed simultaneously, leading to a response conflict in conflict problems but not in no-conflict problems. Therefore, N2 amplitude will be greater in conflict problems than no-conflict problems if the parallel models are correct.

The amplitude of P3 is determined by the allocation of attentional resources when working memory is updated (Donchin & Coles, 1988). This occurs when the mental representation of the current context is revised (Donchin, 1981). In this study, the representation is an interpretation of the conclusion. If evaluation based only on belief occurs first, as predicted by serial models, then there will be a difference in P3 between believable conclusions that are expected and unbelievable conclusions that are not expected. There will be no effect of logical validity as logic is not predicted to influence the conclusion at this stage. If evaluation is based on belief and logic at the same time, as predicted by parallel models, there will be a difference between no-conflict problems that are expected as both logic and belief-based processes lead to the same interpretation, and conflict problems in which the final word is unexpected either logically or compared to prior beliefs.

The latency of P3 reflects stimulus evaluation time (e.g., Kutas, McCarthy, & Donchin, 1977; Magliero, Bashore, Coles, & Donchin, 1984). In this study, the stimuli are conclusions that are evaluated as valid or invalid. If evaluation based on belief occurs first, then the latency of P3 will be determined by belief and not by logic. If evaluation is based on belief and logic at the same time, then the latency of P3 will be determined by the interaction of logic and belief.

Method

Participants

Eighteen psychology students (9 females, 9 males) took part in the experiment in return for an honorarium. Five participants were excluded due to a high number of errors resulting in insufficient trials for analysis. All spoke English as a first language, were right-handed, and had normal or normal-corrected vision.

Materials

Participants were presented with 192 relational reasoning problems, based on the reasoning task of Roberts and Sykes (2003), in which the logic (valid/invalid) and believability (believable/unbelievable) of the conclusions were manipulated in a full factorial design. Examples are presented in Table 1. The problems consisted of three premises and a conclusion. Four parallel versions of each problem were created. The logical validity of the conclusion was manipulated by changing the premises so that the same conclusion was either valid or invalid. A nonsense term was included so that this manipulation did not require premises that conflict with prior knowledge that can be difficult to integrate into a coherent representation (Klauer, Musch, & Naumer, 2000). The believability of the conclusion was manipulated by reversing the elements in the conclusion.

Procedure

Participants silently read the premises and conclusion and then responded either that the conclusion was valid or invalid. They were instructed to assume that the premises were true and to accept the conclusion as valid only if it followed necessarily from the premises. Figure 1 presents a sample trial sequence. After initial instructions to blink and a fixation dot, the premises were presented on a computer monitor, one at a time for 3 s each, with a fixation dot presented between each for 500 ms. The conclusion was then presented with the final word omitted for 2,000 ms. Then, the final word of the conclusion was presented—only at this point was it possible to evaluate whether the conclusion was valid or not. After a further 2,500 ms, the letters V and I appeared on either side of the screen. Participants had two response buttons, one for each hand. If they thought the conclusion was valid, they pressed the response button that corresponded to the side of the screen on which the letter V was presented. If they thought the conclusion was invalid, they pressed the response button that corresponded to the side of the screen on which the letter I was presented. The side on which the letters V and I appeared varied randomly on each trial so that participants could not start preparing their motor response until they appeared. This ensured that the electroencephalogram (EEG) recording between the presentation of the final word of the conclusion and the presentation of response options measured only reasoning and not response execution.

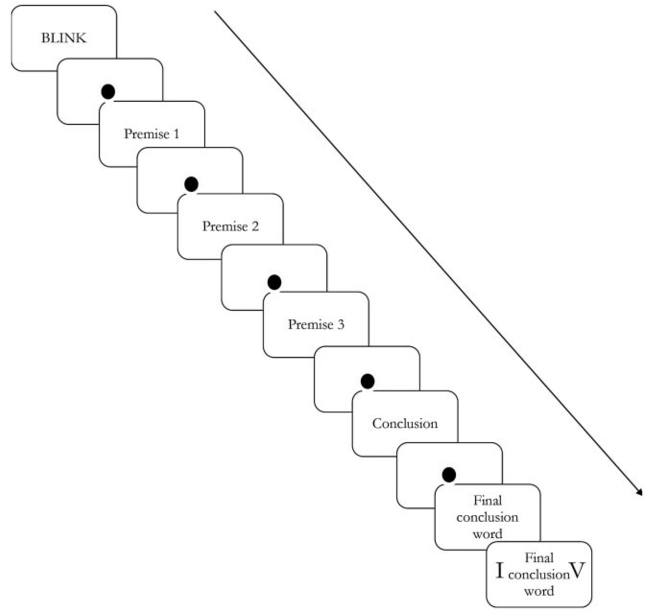

Figure 1. Illustration of the experimental procedure.

EEG Recording and Analysis

EEG activity was recorded from 32 Ag/AgCl scalp electrodes located according to the 10-20 system, with 4 additional electrooculogram (EOG) electrodes placed above and below the right eye and at the outer canthi. Data were recorded in DC mode at 500 Hz using an average reference. Impedance was kept below 5 kΩ. Data were rereferenced offline to linked earlobes and lowpass filtered at 40 Hz (24 dB/octave). Trials containing eye blinks and other eye movement artifacts were rejected when horizontal EOG exceeded ± 40 μV and vertical EOG exceeded ± 80 μV. The data were then segmented into 900-ms epochs, including a 100-ms prestimulus baseline relative to the onset of the final conclusion word. The epochs were baseline corrected using the mean prestimulus voltage in the 100-ms prestimulus period. Average ERPs for correct trials were calculated for each condition separately. The mean number of trials for each condition was: believable valid, 33.5; unbelievable valid, 30.2; believable invalid, 24.1; unbelievable invalid, 31.2. The number of trials analyzed differed across conditions, $F(3,36) = 7.73$, $p < .001$, $\eta^2 = 0.39$. As only logically correct trials were included in the analysis, these differences reflect the predicted pattern of error rate that is necessary to demonstrate the characteristic behavioral effect of belief bias. A supplementary analysis below tests a subgroup of participants with equal trial numbers across conditions and replicates the most theoretically important differences between conditions.

Amplitudes were measured as the mean voltage in a given measurement window. Mean amplitudes for the N2 were calculated at electrode site Fz, FCz, and Cz in the time range 250–370 ms and for the P3 component at electrodes Fz, Cz, and Pz in the time range 300–500 ms. The N2 component was defined as the largest negative-going peak in the time range 200–400 ms. The P3 component was defined as the largest positive-going peak in the time range 300–700 ms. Peak latency of N2 and P3 were calculated as

Table 2. *Mean Percentage of Conclusions Accepted as a Function of Logic and Belief*

	Believable	Unbelievable
Valid	82.53 (9.14)	72.60 (15.16)
Invalid	43.59 (28.16)	19.23 (15.84)

Note. Standard deviations are in parentheses.

the time from stimulus onset to the maximum amplitude of each peak, respectively. All EEG measures were analyzed with a 2 Belief × 2 Logic × 3 Electrode repeated measures analysis of variance (ANOVA). The Greenhouse-Geisser correction was applied when the assumption of sphericity was violated, and uncorrected degrees of freedom, the corrected p value, and epsilon are reported (Jennings & Wood, 1976). Post hoc analyses were conducted using Bonferroni corrected multiple t tests.

Results

Behavioral Results

Table 2 presents the behavioral results for each condition. Valid conclusions were accepted more frequently than invalid conclusions, $F(1,12) = 37.81$, $p < .001$, $\eta_p^2 = .76$. Believable conclusions were accepted more frequently than unbelievable conclusions, $F(1,12) = 13.29$, $p = .003$, $\eta_p^2 = .53$. There was a significant interaction of logic and belief, $F(1,2) = 8.26$, $p = .014$, $\eta_p^2 = .41$. Post hoc analyses showed that there was a marginally nonsignificant difference between believable valid and unbelievable valid conclusions, $t(12) = 2.53$, $p = .027$, with a trend towards accepting more believable conclusions. Believable invalid conclusions were accepted significantly more often than unbelievable invalid conclusions, $t(12) = 3.79$, $p = .003$. The effect of belief on conclusion acceptance was greater for invalid than valid conclusions. This pattern of results replicates the typical findings in studies of belief bias (e.g., Evans, Barston, & Pollard, 1983; Roberts & Sykes, 2003).

ERP Amplitudes

Figure 2 depicts the grand average ERP waveforms and topographic amplitude plots.

N2 amplitude. N2 amplitude was more negative in frontal than central electrode sites, $F(2,24) = 8.02$, $p = .008$, $\varepsilon = .65$, $\eta_p^2 = .40$, with the amplitude at Fz more negative than at Cz, $t(12) = 3.05$, $p = .01$, but there were no significant differences between other electrodes, $t(12) < 2.6$, $p > .02$ (both comparisons). There was no significant effect of logic, $F(1,12) = 1.98$, $p = .185$, $\eta_p^2 = .14$, or belief, $F(1,12) = 0.53$, $p = .481$, $\eta_p^2 = .04$. There was a significant interaction between logic and belief, $F(1,12) = 7.4$, $p = .019$, $\eta_p^2 = .38$. As there was no interaction of belief or logic with electrode site, $F(2,24) < 0.5$, $p > .7$, $\varepsilon = .65$, $\eta_p^2 < .02$ (both comparisons), mean amplitude was pooled across sites for the post hoc tests. There was no significant difference between believable valid and unbelievable valid conclusions, $t(12) = -1.21$, $p = .25$. The difference between believable invalid and unbelievable invalid conclusions was marginally nonsignificant, $t(12) = 2.52$, $p = .027$, with a trend towards greater negativity for unbelievable conclusions.

P3 amplitude. P3 amplitude increased from frontal to parietal electrode sites, $F(2,24) = 24.07$, $p < .001$, $\varepsilon = .94$, $\eta_p^2 = .67$, with the largest amplitude at Pz, followed by Cz and then Fz (all $t(12) > 3.5$, $p \leq .004$). There was no significant effect of logic, $F(1,12) = 0.83$, $p = .380$, $\eta_p^2 = .07$, and no effect of belief, $F(1,12) = 0.001$, $p = .977$, $\eta_p^2 = .00$. There was a significant interaction between logic and belief, $F(1,12) = 13.61$, $p = .003$, $\eta_p^2 = .53$. As there was no interaction of belief or logic with electrode site, $F(2,24) < 1.3$, $p > .25$, $\varepsilon = .94$, $\eta_p^2 < .1$ (both comparisons), mean amplitude for belief and logic was pooled across sites for the post hoc tests. Mean P3 amplitude was greater for unbelievable valid than believable valid conclusions, but this difference fell marginally short of significance after Bonferroni correction, $t(12) = -2.55$, $p = .026$. The pattern was reversed for invalid conclusions: P3 amplitude was larger for believable invalid than unbelievable invalid conclusions, $t(12) = 3.03$, $p = .011$. P3 amplitude was larger in problems where belief and logic conflict than in no-conflict problems. There was a three-way interaction of electrode site, logic, and belief, $F(2,24) = 7.37$, $p = .008$, $\varepsilon = .94$, $\eta_p^2 = .38$, indicating that, following Bonferroni correction, the interaction of logic and belief was present at Cz, $F(1,12) = 13.16$, $p = .003$, $\eta_p^2 = .52$, and Pz, $F(1,12) = 20.21$, $p = .001$, $\eta_p^2 = .63$, but not at Fz, $F(1,12) = 4.42$, $p = .057$, $\eta_p^2 = .27$.

ERP Latencies

Figure 2 depicts the grand average ERP waveforms and topographic amplitude plots.

N2 latency. N2 latency increased from central to frontal electrode sites, $F(2,24) = 13.56$, $p = .002$, $\varepsilon = .61$, $\eta_p^2 = .53$, with the shortest latency at Cz, followed by FCz and then Fz (all $t(12) > 3.1$, $p \leq .002$). No other main effects or interactions were significant (all $F < 1.1$, $p > 0.3$, $\eta_p^2 < .08$.)

P3 latency. P3 latency increased from frontal to parietal electrode sites, $F(2,24) = 7.54$, $p = .006$, $\varepsilon = .82$, $\eta_p^2 = .39$, with Fz shorter than Pz, $t(12) = 3.63$, $p = .003$. Fz was not shorter than Cz, although this comparison approached significance, $t(12) = 2.37$, $p = .036$. Cz and Pz did not differ, $t(12) = 1.29$, $p = .222$. There was a significant interaction between electrode site and belief, $F(2,24) = 4.32$, $p = .025$, $\varepsilon = .82$, $\eta_p^2 = .27$. P3 latency for believable conclusions were shorter than unbelievable conclusions at Fz, $t(12) = 3.12$, $p = .009$, but there were no differences at Cz or Pz, $t(12) < 1.2$, $p > .25$ (both comparisons). The interaction of belief and logic was nonsignificant, $F(1,12) = 2.41$, $p = .147$, $\eta_p^2 = .17$. No other main effects or interactions were significant (all $F < 0.5$, $p > 0.6$, $\eta_p^2 < .04$.)

Supplementary Analyses

A characteristic behavioral belief bias effect was found in which belief influenced reasoning resulting in different logical error rates across conditions. A consequence of this is that the number of trials in each condition varied because only logically correct trials were analyzed, and this can bias ERPs. Our main findings were differences in amplitude across condition. We analyzed this using mean amplitude, which is not biased by number of trials, unlike peak amplitude, and so the differences on this measure are unlikely to be the result of an artifact (Luck, 2005). Nonetheless, to test this possibility, we replicated the analyses using a subgroup of eight participants who did not differ significantly in the number of trials

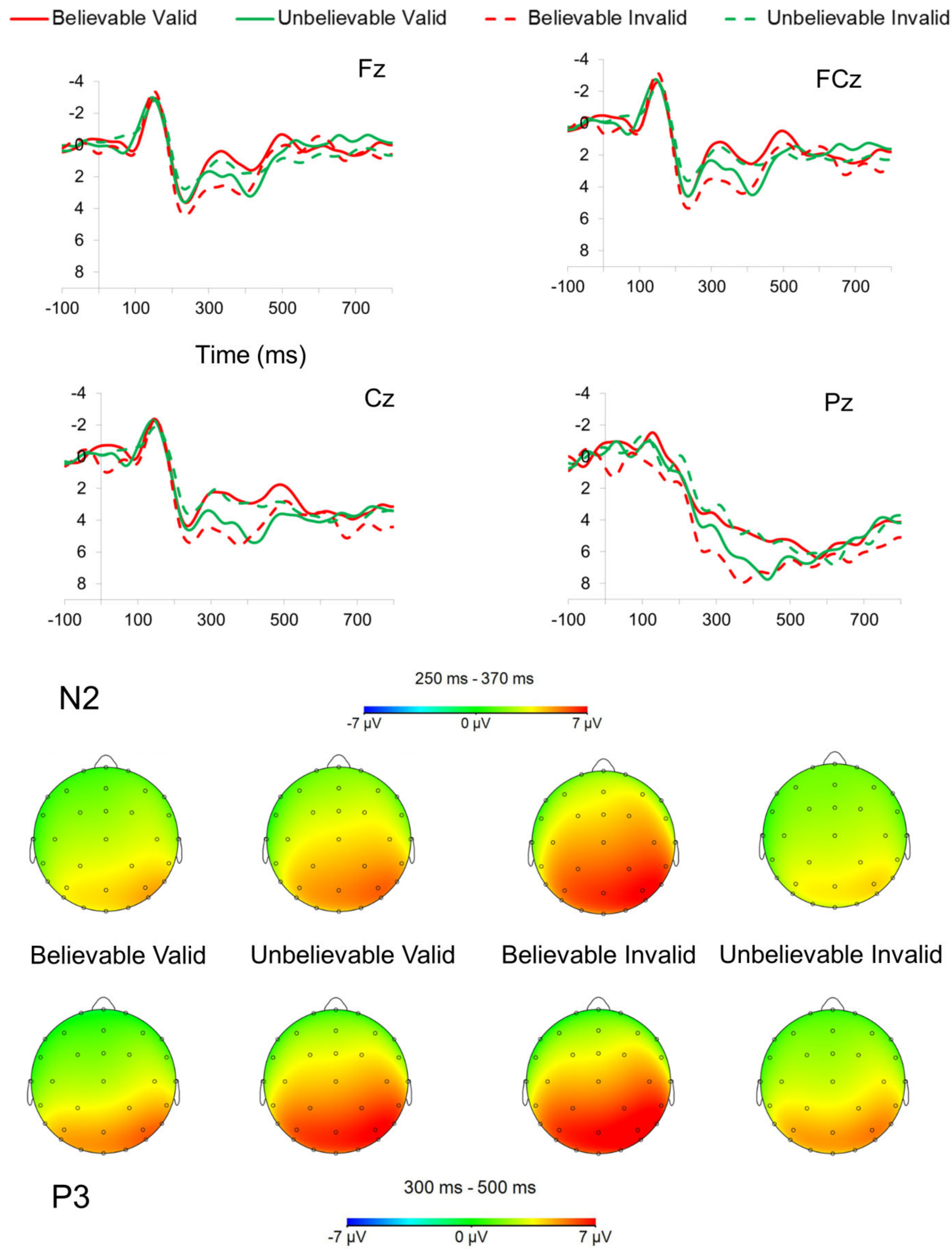

Figure 2. Stimulus-locked ERP waveforms at Fz, FCz, Cz, and Pz as a function of logic and belief. Scalp topography for the 250–370 ms time window as a function of logic and belief (N2). Scalp topography for the 300–500 ms time window as a function of logic and belief (P3).

per condition, $F(3,18) = 2.91$, $p = .059$, $\eta_p^2 = .29$. The mean number of trials per condition was: believable valid, 35.25; unbelievable valid, 32.38; believable invalid, 30.38; unbelievable invalid, 35.75.

The significant differences in N2 amplitude found in the main analysis were not replicated but similar trends were found that approached significance. The comparison of electrode sites was not significant, $F(2,14) = 4.20$, $p = .069$, $\varepsilon = .60$, $\eta_p^2 = .38$. The interaction of belief and logic fell short of significance, $F(1,7) = 5.16$, $p = .057$, $\eta_p^2 = .42$. The remaining comparisons were also not significant, $F < 1.1$, $p > .5$, $\eta_p^2 < .13$ (all comparisons).

The most theoretically important differences in P3 amplitude found in the main analysis were replicated in the subgroup analysis. P3 amplitude increased from frontal to parietal electrode sites, $F(2,14) = 8.58$, $p = .005$, $\varepsilon = .91$, $\eta_p^2 = .55$. Pz amplitude was larger than Fz, $t(7) = 3.61$, $p = .009$, but the other comparisons were not significant: both $t(7) < 2.3$, $p > 0.05$. There was no main effect of logic, $F(1,7) = .12$, $p = .738$, $\eta_p^2 = .02$, or belief, $F(1,7) = 2.41$, $p = .164$, $\eta_p^2 = .26$. There was an interaction of logic and belief, $F(1,7) = 7.63$, $p = .028$, $\eta_p^2 = .52$. Mean P3 amplitude was greater for unbelievable valid than believable valid conclusions, $t(7) = 2.92$, $p = .022$, but for invalid conclusions there was no significant difference between believable invalid and unbelievable invalid conclusions, $t(7) = 1.71$, $p = .13$, although the difference was in the same direction as in the main analysis. The remaining comparisons were not significant, $F < 1.3$, $p > .3$, $\eta_p^2 \leq .15$.

Our latency measure was peak latency, which can be biased by differences in trial numbers. Unlike the main analysis, N2 latency did not differ across electrode site, $F(2,14) = 4.08$, $p = .065$, $\eta_p^2 = .37$, although the trend was in a similar direction. No other comparisons were significant, $F < 2.4$, $p > .1$, $\eta_p^2 < .26$ (all comparisons). There were no significant effects of P3 latency, $F < 2.4$, $p > .1$, $\eta_p^2 \leq .25$ (all comparisons).

Overall, the supplementary analysis confirmed the main P3 amplitude findings—specifically the absence of a main effect of logic or belief but the presence of a crossover interaction of these two factors. P3 amplitude is larger in conflict than in no-conflict problems. In combination, the main and supplementary analyses indicate that this is a reliable effect. N2 amplitude findings and N2 and P3 latency findings were not supported and are less reliable.

Discussion

Dual process theories propose that both heuristic and analytic processes are used during reasoning. Serial models claim that heuristic processes precede analytic processes. Parallel models claim that they occur at the same time. The aim of this study was to compare these two models. A typical belief bias paradigm was used with the addition of ERP recordings to identify when different reasoning processes occur. The results support the predictions of parallel models.

The behavioral findings were as expected. Participants were more likely to accept logical conclusions and more likely to accept believable conclusions. There was also an interaction between these two factors. Participants were less likely to accept unbelievable conclusions than believable conclusions when the problems were invalid than when the problems were valid. This is the characteristic pattern of results in belief bias experiments (e.g., Evans, Barston, & Pollard, 1983; Roberts & Sykes, 2003).

P3 and N2 Amplitudes Were Greater When Logic and Belief Conflict

The most reliable effect, present in both the main and supplementary analyses, was the effect of the interaction of logic and belief on P3 amplitude without a main effect of either factor. P3 was similar for conflict problems (i.e., believable invalid and unbelievable valid problems) and similar for no-conflict problems, (i.e., believable valid and unbelievable invalid problems), but greater in conflict problems than no-conflict problems. The difference between conflict and no-conflict problems requires an evaluation of both logic and belief. Hence, problems were not evaluated initially on the basis of belief or logic alone as predicted by serial models; it is the combination of these factors occurring simultaneously that influences P3. The interaction of logic and belief supports parallel models.

A similar interaction was found in N2 amplitude, but it was not in the expected direction. N2 was larger for no-conflict problems than for conflict problems, whereas it had been predicted that N2 would be larger for conflict problems as these require a conflicting response to be inhibited (e.g., De Neys, Novitskiy et al., 2010). The time window for N2 was chosen based on peak latency of this component, but it does overlap with the P3 window. This novel task may not have adequately elicited N2, also reflected in the absence of N2 effects in the supplementary analysis, and a further study is required to isolate the N2 component and establish how it is affected by belief bias in reasoning.

P3 and N2 Latencies Were Not Influenced by Logic or Belief

There was no overall effect of belief or logic on P3 latency. P3 latency is associated with the time taken to evaluate a stimulus (Kutas et al., 1977; Magliero et al., 1984). This indicates that the time taken to evaluate the conclusion is the same for problems that differ either in belief or logic. This supports parallel models in which these processes occur at the same time. The exception to this finding is an interaction of belief with electrode site. P3 latency was shorter for believable than unbelievable conclusions at Fz. However, this result is not reliable. Peak latencies are sensitive to differences in trial number across the condition that can lead to artifactual results. There were unequal trial numbers in the main analysis, and this finding was not replicated in the supplementary analyses on participants with similar trial numbers. Future studies should seek to replicate this finding with larger trial numbers. There were no differences between conditions in the latencies of N2.

Theoretical Implications for Dual Process Theories

The findings provide support for dual process theories of reasoning in general. Studies of dual process theory have been criticized for failing to rule out alternative single system explanations of the data. For example, Keren and Schul (2009) claim that there has so far been no direct evidence that people hold two interpretations (i.e., belief and logic) concurrently. Our ERP findings do demonstrate that these influences occur simultaneously, obviating this criticism and supporting dual process accounts.

The findings of this study are compatible with parallel models. One recent example is Handley et al's (2011) parallel competitive account. This proposes that both belief and logic may be processed by either heuristic or analytic processes. These two processes operate simultaneously, and each cues a response. In this task, we assume that the heuristic response is a belief-based evaluation, and the logical response is an analytic evaluation. These processes operate in parallel to influence the evaluation of the conclusion.

Another parallel model is De Neys's logical intuition model (De Neys, 2012). This theory is unusual in that both the logical and belief-based processes are "intuitive"; that is, they are automatic evaluations that do not require analytic thinking. The findings of this study are consistent with this theory, but their interpretation according to this theory is that both belief-based and logical processes are intuitive. The analytic process occurs later. While this study suggests that logic and belief are evaluated in parallel, further work is required to distinguish between different parallel models such as those in which heuristic and analytic processes occur at the same time (e.g., Handley et al., 2011; Sloman, 1996) and De Neys's (2012) model in which nonanalytic intuitions can be both logical and belief based.

The findings are not consistent with the main serial model, the default-interventionist account (Evans, 2006). This explanation of belief bias is that an initial default solution is provided using a heuristic process based on the believability of the conclusion, but an analytic process may intervene and override this response. This model does not explain the influence of logic on conclusion evaluation that was found in P3 as it predicts that only belief influences conclusion evaluation at that stage. Some support was found for this model in the analysis of peak latencies, which were shorter for believable than unbelievable conclusions at Fz but were not influenced by logic, indicating a faster evaluation of belief than logic. However, this result was not replicated in the supplementary analysis and is not reliable.

It is possible that the design of this study in which the premises were presented serially prior to the conclusion encouraged "forward" premise-driven reasoning from the premises to the conclusion rather than "backwards" conclusion-driven reasoning[1] (Morley, Evans, & Handley, 2004). As influential default-interventionist theories of belief bias such as selective processing (Evans, 2007; Stupple et al., 2011) suppose that reasoning begins with an evaluation of the conclusion, this procedure may not be a good test of that theory. However, this study does not seek to compare premise-driven or conclusion-driven accounts. Our aim is to test more generally the properties of parallel and serial models. Our findings show that, at the time when the conclusion is presented, there is both an influence of logic and of belief. Therefore, we show at least one situation in which there is a parallel influence of belief and logic on conclusion evaluation. Future research would be required to further investigate the generality of this finding in premise-driven and conclusion-driven reasoning.

Our explanation of the parallel effects of logic and belief is that as a person reads a problem they build a representation of it. The representation might be a mental model of a logical possibility, a believable possibility, or both. These representations create a semantic context that influences the response to later words. Words that do not fit the semantic context are unexpected and require a change to the representation that is updated in working memory, eliciting a larger P3 amplitude (Polich, 2007). This is the case in conflict problems because logical and belief-based interpretations are incompatible, and the new word cannot be consistent with both of them. Words that do fit the representation do not require it to be changed and elicit a smaller P3. This is the case in no-conflict problems because new words are compatible with both logic and belief-based interpretations.

Conclusion

This study found greater P3 amplitude when evaluating conflict problems compared to no-conflict problems. The detection of a conflict between a belief-based response and a logical response is consistent with other research (e.g., De Neys, Moyens, & Vansteenwegen, 2010; De Neys & Glumicic, 2008; De Neys et al., 2008). This study differs in that it shows *when* the two processes occur: they occur relatively early and in parallel. This provides support for dual process theories of reasoning in which belief-based and logical evaluations occur in parallel but not theories in which belief-based heuristic evaluations precede logical analysis.

1. We thank an anonymous reviewer for this suggestion.

References

Andrews, G. (2010). Belief-based and analytic processing in transitive inference depends on premise integration difficulty. *Memory & Cognition*, *38*, 928–940. doi: 10.3758/MC.38.7.928

Banks, A. P. (2013). The influence of activation level on belief bias in relational reasoning. *Cognitive Science*, *37*, 544–577. doi: 10.1111/cogs.12017

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

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299. doi: 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*, 208–216. doi: 10.3758/CABN.10.2.208

De Neys, W., Novitskiy, N., Ramautar, J., & Wagemans, J. (2010). What makes a good reasoner?: Brain potentials and heuristic bias susceptibility. In S. Ohlsson, & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1020–1025). Austin, TX: Cognitive Science Society.

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

Donchin, E. (1981). Surprise! . . . Surprise? *Psychophysiology*, *28*, 177–184. doi: 10.1111/j.1469-8986.1981.tb01815.x

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral Brain Science*, *11*, 357–374. doi:10.1017/S0140525X00058027

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

Evans J. S. B. T. (2006). The heuristic-analytic theory of reasoning: extension and evaluation. *Psychonomic Bulletin & Review*, *13*, 378–395. doi: 10.3758/BF03193858

Evans, J. S. B. T. (2007). On the resolution of conflict in dual-process theories of reasoning. *Thinking & Reasoning*, *13*, 321–329. doi: 10.1080/13546780601008825

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. doi: 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*, 295–306.

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*, 382–389. doi: 10.1080/13546780542000005

Evans, J. S. B. T., Handley, S. J., & Bacon, A. M. (2009). Reasoning under time pressure: A study of causal conditional inference. *Experimental Psychology*, *56*, 77–83. doi: 10.1027/1618-3169.56.2.77

Evans, J. S. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.

Folstein, J. R., & van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*, 152–170. doi: 10.1111/j.1469-8986.2007.00602.x

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*, 28–43. doi: 10.1037/a0021098

James, W. (1950). *The principles of psychology*. New York, NY: Dover. (Original work published 1890).

Jennings, J. R., & Wood, C. C. (1976). The ε-adjustment procedure for repeated-measures analysis of variance. *Psychophysiology*, *13*, 277–278.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge, UK: Cambridge University Press.

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science*, *4*, 533–550. doi: 10.1111/j.1745-6924.2009.01164.x

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*, 852–884. doi: 10.1037/0033-295X.107.4.852

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*, 792–795.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luo, J., Liu, X., Stupple, E. N., Zhang, E., Xiao, X., Lei, J., . . . Zhang, Q. (2013). Cognitive control in belief-laden reasoning during conclusion processing: An ERP study. *International Journal of Psychology*, *48*, 224–231. doi: 10.1080/00207594.2012.677539

Luo, J., Yuan, J., Qiu, J., Zhang, Q., Zhong, J., & Zhangcui, H. (2008). Neural correlates of the belief-bias effect in syllogistic reasoning: An event-related potential study. *NeuroReport*, *19*, 1073–1078. doi: 10.1097/WNR.0b013e3283052fe1

Magliero, A., Bashore, T. R., Coles, M. G. H., & Donchin, E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology*, *21*, 171–186. doi: 10.1111/j.1469-8986.1984.tb00201.x

Morley, N. J., Evans, J. S. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, *57A*, 666–692. doi: 10.1080/02724980343000440

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Roberts, M. J., & Sykes, E. D. A. (2003). Belief bias and relational reasoning. *Quarterly Journal of Experimental Psychology*, *56A*, 131–154. doi: 10.1080/02724980244000233

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. doi: 10.1037/0033-2909.119.1.3

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–726. doi: 10.1017/S0140525X00003435

Stupple, E. J. N., Ball, L. J., Evans, J. S. B. T., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology*, *23*, 931–941. doi: 10.1080/20445911.2011.589381

(Received August 23, 2012; Accepted September 21, 2013)