

Reasoning as we read: Establishing the probability of causal conditionals

Matthew Haigh · Andrew J. Stewart · Louise Connell

Published online: 5 September 2012
© Psychonomic Society, Inc. 2012

Abstract Indicative conditionals of the form *if p then q* (e.g., *if student tuition fees rise, then applications for university places will fall*) invite consideration of a hypothetical event (e.g., tuition fees rising) and of one of its possible consequences (e.g., applications falling). Since a rise in tuition fees is an uncertain event with equally uncertain consequences, a reader may believe the statement to a greater or lesser extent. As a conditional is read, the earliest point at which this probabilistic evaluation can take place is as the consequent clause is wrapped up (e.g., as the critical word *fall* is read in the example above). Wrap-up processing occurs at the end of the clause, as it is evaluated and integrated into the evolving discourse representation. Five sources of probability may plausibly influence the evaluation of a conditional as it is wrapped up; these are $P(p)$, $P(q)$, $P(pq)$, $P(q|p)$, and $P(\text{not-}p \text{ or } q)$. A total of 128 conditionals were constructed, with these probabilities calculated for each item in a pretest. The conditionals were then embedded in vignettes and read by 36 participants on a word-by-word basis. Using linear mixed-effects modeling, we found that wrap-up reading times were predicted by pretest ratings of $P(p)$ and $P(q|p)$. There was no influence of $P(q)$, $P(pq)$, or $P(\text{not-}p \text{ or } q)$ on wrap-up reading times. Our findings are consistent with the suppositional theory of conditionals proposed by Evans and Over (2004) but do not support the mental-models theory advanced by Johnson-Laird and Byrne (2002).

Keywords Reasoning · Psycholinguistics · Mental models · Conditionals · if · Supposition

Indicative conditionals of the form *if p then q* invite the consideration of a hypothetical event (p) and one of its possible consequences (q). For instance, a newspaper opinion piece that asserts *if student tuition fees rise, then applications for university places will fall* encourages the reader to mentally entertain a possible state of the world in which the number of university applications falls following a rise in student tuition fees. Since a rise in tuition fees is an uncertain future event with equally uncertain consequences, a reader may believe the statement to a greater or lesser extent. This subjective degree of belief can be quantified as the probability of the conditional, or $P(\text{if } p \text{ then } q)$.

The ability to rapidly evaluate the probability of a conditional describing a hypothetical event is central to everyday reasoning and decision making (see Evans & Over, 2004). However, no consensus exists about how people subjectively establish $P(\text{if } p \text{ then } q)$. Some have argued that this judgment is equivalent to the subjective conditional probability [$P(q|p)$] of the consequent event given the antecedent event (e.g., the probability that student applications will fall given a rise in tuition fees; Evans & Over, 2004). Others have suggested that people base their belief initially on the subjective conjunctive probability [$P(pq)$] of the antecedent and consequent events occurring together (e.g., the probability that both tuition fees will rise *and* applications will fall) but can, in some cases, also arrive at a conclusion by thinking about all of the possibilities in which the conditional is true [i.e., $P(\text{not-}p \text{ or } q)$] (Johnson-Laird & Byrne, 1991, 2002). In this article, we will examine the processing loads associated with $P(p)$, $P(q)$, $P(pq)$, $P(q|p)$, and $P(\text{not-}p \text{ or } q)$ to determine which probabilities readers use to rapidly guide their evaluations of a conditional during comprehension.

Within the conditional-reasoning literature, the Ramsey test is an influential perspective that describes a mechanism for engaging in hypothetical thought. Ramsey proposed that people judge their belief in conditionals of the form *if p then q* by

Electronic supplementary material The online version of this article (doi:10.3758/s13421-012-0250-0) contains supplementary material, which is available to authorized users.

M. Haigh (✉) · A. J. Stewart · L. Connell
School of Psychological Sciences, University of Manchester,
Manchester M13 9PL, UK
e-mail: matthew.haigh@manchester.ac.uk

A. J. Stewart
e-mail: andrew.stewart@manchester.ac.uk

L. Connell
e-mail: louise.connell@manchester.ac.uk

“. . . adding p hypothetically to their stock of knowledge and arguing on that basis about q . . . [fixing] their degrees of belief in q given p . . .” (Ramsey, 1931/1990, p. 247). The Ramsey test has been formalized by psychologists in the field of human reasoning as the suppositional theory of *if* (Evans & Over, 2004). The suppositional theory proposes that people use epistemic mental models to evaluate their degree of belief in a conditional. This degree of belief is established by making a probability judgment about the extent to which they believe that the consequent event will occur within a hypothetical world in which the antecedent is true (e.g., following the example above, this would be the subjective probability of student applications falling given a rise in tuition fees). This probability judgment is known as the subjective conditional probability, or $P(q|p)$, and has been shown to play a central role in how conditionals are ultimately interpreted (e.g., Oberauer & Wilhelm, 2003).

An alternative perspective is based on the idea that people represent conditional information using semantic (rather than epistemic) mental models (Johnson-Laird & Byrne, 1991, 2002). The mental-models theory proposes that people mentally represent the truth-verifiable possibilities asserted by a conditional (rather than the possibilities in which p holds, as the suppositional theory claims). For an indicative conditional of the form *if p then q* , these possibilities are the truth table rows that make the conditional true (see Table 1).

An important feature of the model theory is that the initial representation of a conditional statement only makes explicit the p & q case, with the other possibilities being implicit until they are required (as denoted below by the ellipsis).

p & q

. . .

If required, this initial model can then be fleshed out to represent all of the states of the world in which the statement is true. This makes the fully fleshed-out model equivalent to the truth-functional material conditional of propositional logic, which is always true in cases containing not- p or q .

p & q

not- p & q

not- p & not- q

Table 1 Truth table for *if p then q*

p	q	If p then q
p	q	True
p	not- q	False
not- p	q	True
not- p	not- q	True

In terms of establishing degrees of belief in a conditional statement, it has been argued that these mental models can be used to determine $P(\text{if } p \text{ then } q)$ (Giroto & Johnson-Laird, 2004; Johnson-Laird, Legrenzi, Giroto, Legrenzi, & Caverni, 1999). This can be achieved in two ways. Firstly, because people will often fail to flesh out their initial mental model (e.g., due to working memory limitations), they will simply base their belief in a conditional on the probability of the initial model—that is, $P(pq)$ (e.g., the probability that both tuition fees will rise and applications will fall). Alternatively, if this initial model is successfully fleshed out, belief in the conditional can be calculated by summing the probabilities of the models in which the statement is true (Johnson-Laird et al., 1999). The probability of this fully fleshed-out mental model is equivalent to the probability of the material conditional [i.e., $P(\text{not-}p \text{ or } q)$].

To examine how participants judge the probability of conditionals, Evans, Handley, and Over (2003) presented abstract conditional statements and the associated probability distributions (e.g., if the card is yellow, then it has a circle printed on it). They attempted to reveal which of three probabilities participants would use to establish their degree of belief in a conditional statement [i.e., $P(\text{if } p \text{ then } q)$]. These probabilities were $P(q|p)$, $P(pq)$, and $P(\text{not-}p \text{ or } q)$ (i.e., the probability of the material conditional). They found no evidence that people base their belief on the probability of the material conditional, but rather found that participants fell into two groups. One group based their belief on $P(q|p)$, while the other, slightly smaller group based their belief on $P(pq)$ (see also Oberauer & Wilhelm, 2003; Politzer, Over, & Baratgin, 2010, for similar findings). It has since been shown that adults who initially judge a conditional as $P(pq)$ tend to switch to a $P(q|p)$ interpretation as more and more trials are presented (Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011). The influence of $P(pq)$ has been attributed to a form of shallow processing (Evans et al., 2003) and also to individual differences in cognitive ability (Evans, Neilens, & Over, 2008); however, this effect has not been consistently replicated in the literature (Evans, Handley, Neilens, & Over, 2007).

Only recently has attention turned to how the comprehension of everyday causal conditionals might be influenced by our real-world knowledge. Over, Hadjichristidis, Evans, Handley, and Sloman (2007) examined the probability of everyday conditional statements (e.g., if the cost of petrol increases, then traffic congestion will improve) by asking participants to assign probabilities to the truth table conjunctions p & q , p & not- q , not- p & q , and not- p & not- q . From these ratings, Over et al. calculated three statistically independent predictors that could be used to determine whether people base their belief in a conditional on the conditional probability, the conjunctive probability, or the probability of the fully fleshed-out material conditional. Analyses revealed that $P(q|p)$ was a strong predictor of subjective ratings of $P(\text{if } p \text{ then } q)$, thus providing support

for a conditional-probability hypothesis. There was some weak evidence for a conjunctive-probability interpretation, and no support for the fully fleshed-out material conditional hypothesis.

Experiment: Reasoning as we read

While there is evidence that $P(q|p)$ and, to a lesser extent, $P(pq)$ inform belief in a conditional statement, little is known about the mechanisms that guide the fast-acting comprehension processes required to understand conditionals as they are processed in real time. Traditional measures of conditional reasoning (e.g., the deduction paradigm) rely on inferences or decisions that people make or endorse following a conditional statement. These tasks typically require sustained analytical processing, which is in contrast to the rapid and intuitive comprehension associated with conditional statements in everyday discourse. As a result, these offline techniques can only provide data based on the *ultimate* representation of a conditional. By focusing only on this final, fully formed representation, a number of distinct, incremental processes necessary for comprehension of a conditional statement may be overlooked.

In the experiment reported below, we employed a word-by-word self-paced reading paradigm to examine comprehension processes as the interpretation of a conditional is built in real time. A time-course approach has recently provided new insights into both the processing of conditional statements themselves (Espino, Santamaria, & Byrne, 2009; Haigh & Stewart, 2011; Stewart, Haigh, & Kidd, 2009) and the processing of information following a conditional statement (e.g., de Vega, Urrutia, & Rizzo, 2007; Ferguson, 2012; Ferguson & Sanford, 2008; Haigh, Stewart, Wood, & Connell, 2011; Rader & Sloutsky, 2002).

Specifically, we examined the influence of $P(p)$, $P(q)$, $P(pq)$, $P(q|p)$, and $P(\text{not-}p \text{ or } q)$ on reading times to everyday causal conditionals that were embedded in vignettes. The dependent measure of interest was the reading time for the critical region of text at the end of the consequent clause (i.e., the earliest point at which the conditional could be evaluated as a whole). Reading times for this region capture *wrap-up* processing, which occurs at the end of a clause as information is evaluated and integrated into the evolving discourse representation (Rayner, Kambe, & Duffy, 2000). For our conditionals, the wrap-up region was always the final word of the consequent clause immediately preceding the end of a sentence (e.g., . . . *if student tuition fees rise, then applications for university places will fall*).

Reading times tend to be negatively associated with the subjective plausibility of a clause or sentence, with increased latencies as subjective plausibility decreases

(Rayner, Warren, Juhasz, & Livesedge, 2004). Evidence that ratings of $P(q|p)$ negatively predict reading times to the wrap-up region would therefore provide support for the suppositional theory of Evans and Over (2004), who argued that processes approximating a Ramsey test are engaged to rapidly establish the subjective conditional probability. In contrast, evidence that ratings of $P(pq)$ or $P(\text{not-}p \text{ or } q)$ negatively predict wrap-up reading times would be consistent with the mental-models theory of conditionals developed by Johnson-Laird and Byrne (2002). Specifically, finding that $P(pq)$ predicts reading times would indicate that readers only represent an initial mental model during the early stages of comprehension, whereas an association with $P(\text{not-}p \text{ or } q)$ would indicate that readers rapidly flesh out their mental model.

Method

To ensure sufficient variance in reading times, the 128 conditionals used in our experiment were constructed from clauses that had either high or low subjective probability, as determined by a pretest of the materials (see below for details). These high- and low-probability clauses were fully counterbalanced across items. Because $P(\text{if } p \text{ then } q)$ varies as an interaction of $P(p)$ and $P(q)$, we also ensured that this variable was counterbalanced. For example, a conditional with a high $P(p)$ and high $P(q)$, such as *if the cost of oil rises, then the cost of petrol will rise*, could have an intuitively high $P(\text{if } p \text{ then } q)$. However, another conditional with a similarly high $P(p)$ and high $P(q)$ could have an intuitively low $P(\text{if } p \text{ then } q)$ (e.g., *if more money is spent annually on preventing heart disease, then levels of heart disease will increase*). We counterbalanced all eight possible permutations of high and low $P(p)$, $P(q)$, and $P(\text{if } p \text{ then } q)$ in a $2 \times 2 \times 2$ design (see Table 2).

Pretests

Individual-clause probability task A rating task was carried out to generate conditional statements with antecedents and consequents that had high versus low subjective probabilities of occurring over the next 10 years. A group of 24 students from the University of Manchester were presented with 64 statements (e.g., *Over the next 10 years, student tuition fees will decrease*) and were asked to rate the probability of these events on a scale of 0–100. All probability judgments were collected in early 2009.

The average rating for antecedent statements selected as having high probability ($M = 73$, $SE = 1.21$) was significantly higher than that for antecedent statements selected as having low probability ($M = 27$, $SE = 1.28$), $t_1(23) = 12.73$, $p < .001$; $t_2(31) = 21.57$, $p < .001$. The average rating for

Table 2 Example materials showing all eight permutations of high and low $P(p)$, $P(q)$, and $P(\text{if } p \text{ then } q)$

Example of Items 1–16 (1–64)

Conflicts in the Middle East have destabilised the price of oil subsequently affecting petrol prices in the UK. This has caused a great deal of uncertainty for consumers. Analysts predict that . . .

- (a) . . . if the cost of oil rises, then the cost of petrol will rise. [high $P(p)$, high $P(q)$, high $P(\text{if } p \text{ then } q)$]
- (b) . . . if the cost of oil rises, then the cost of petrol will fall. [high $P(p)$, low $P(q)$, low $P(\text{if } p \text{ then } q)$]
- (c) . . . if the cost of oil falls, then the cost of petrol will rise. [low $P(p)$, high $P(q)$, low $P(\text{if } p \text{ then } q)$]
- (d) . . . if the cost of oil falls, then the cost of petrol will fall. [low $P(p)$, low $P(q)$, high $P(\text{if } p \text{ then } q)$]

Consumers need to be aware of these potential changes.

Example of Items 17–32 (65–128)

The British Heart Foundation continually campaigns to generate funds for research into heart disease. This research could potentially save many thousands of lives. They believe strongly that . . .

- (a) . . . if more money is spent annually on preventing heart disease, then levels of heart disease will increase. [high $P(p)$, high $P(q)$, low $P(\text{if } p \text{ then } q)$]
- (b) . . . if more money is spent annually on preventing heart disease, then levels of heart disease will decrease. [high $P(p)$, low $P(q)$, high $P(\text{if } p \text{ then } q)$]
- (c) . . . if less money is spent annually on preventing heart disease, then levels of heart disease will increase. [low $P(p)$, high $P(q)$, high $P(\text{if } p \text{ then } q)$]
- (d) . . . if less money is spent annually on preventing heart disease, then levels of heart disease will decrease. [low $P(p)$, low $P(q)$, low $P(\text{if } p \text{ then } q)$]

Donations can be made to the British Heart Foundation through their website.

We constructed four versions of 32 items (128 items in total)

consequent statements selected as having high probability ($M = 70$, $SE = 1.42$) was significantly higher than that for consequent statements selected as having low probability ($M = 29$, $SE = 1.55$), $t_1(23) = 11.42$, $p < .001$; $t_2(31) = 15.37$, $p < .001$.

Truth table task (Over et al., 2007) To calculate $P(q|p)$, $P(pq)$, and $P(\text{not-}p \text{ or } q)$ for each conditional, we asked participants to assign probabilities to the truth table conjunctions TT, TF, FT, and FF occurring over the next 10 years. The participants were instructed that these four ratings should sum to 100 (see Over et al., 2007, for a full description of this method). The truth tables were rated by 17 participants. These participants did not take part in the individual-clause rating task. Probabilities were calculated as follows:

$$P(pq) = P(\text{TT})$$

$$P(q|p) = P(\text{TT})/[P(\text{TT}) + P(\text{TF})]$$

$$P(\text{not-}p \text{ or } q) = P(\text{TT}) + P(\text{FT}) + P(\text{FF}).$$

Table 3 Correlations between each of the five probabilities calculated in pretests

	$P(p)$	$P(q)$	$P(pq)$	$P(q p)$	$P(\text{not-}p \text{ or } q)$
$P(p)$	–	.001	.464*	.001	.374*
$P(q)$.001	–	.441*	.479*	–.491*
$P(pq)$.464*	.441*	–	.829*	.421*
$P(q p)$.001	.479*	.829*	–	.285*
$P(\text{not-}p \text{ or } q)$.374*	–.491*	.421*	.285*	–

* $p \leq .001$

The correlations between each of the five probabilities calculated in the pretests are reported in Table 3.

Comprehension task We used the results of our two offline tasks to generate the 128 experimental items examined in the comprehension task.

Participants

A group of 36 volunteers from the University of Manchester population took part in the reading study. All participants were native English speakers and did not have a reading disability. They were each paid £5. Participants who took part in this comprehension task had not taken part in either of the offline ratings tasks.

Materials

A total of 128 experimental vignettes were used in this study. Each vignette consisted of four sentences (see Table 2). The first two sentences provided context, Sentence 3 contained the indicative conditional, and Sentence 4 provided additional contextual information. The full list of experimental vignettes with their associated probabilities can be found in the [Electronic supplementary materials](#). These vignettes were divided into four lists using a repeated measures Latin-square design, with each list containing 32 of the 128 items. Each list also contained 16 filler passages. These filler passages were each four sentences long and did not contain conditionals. All participants saw equal numbers of passages across the eight counterbalanced conditions, and 48 items in total (32 experimental plus 16 filler).

Table 4 Regression weights and confidence intervals (CIs) in linear mixed regression models for each critical region (parameter estimates and p values based on the t statistic)

	Region 1 (Wrap-Up Region)					Region 2 (Spill Over)				
	Estimate	95 % CI Lower	95 % CI Upper	p	r	Estimate	95 % CI Lower	95 % CI Upper	p	r
Intercept	1,163.6	514.3	1,812.8	<.001	–	892.5	428.3	1,356.7	<.001	–
$P(p)$	–4.4	–8.1	–0.7	.019	–.078*	0.74	–1.9	3.4	.589	–.011
$P(q)$	2.0	–2.5	6.7	.376	–.027	1.1	–2.2	4.4	.508	.015
$P(pq)$	5.2	–8.0	18.3	.443	–.131*	1.6	–11.1	8.0	.748	–.101*
$P(q p)$	–10.6	–18.1	–3.1	.006	–.128*	–3.1	–8.6	2.3	.256	–.115*
$P(\text{not-}p \text{ or } q)$	1.0	–6.9	8.9	.802	–.084*	–3.2	–8.9	2.5	.275	–.110*

Correlation coefficients (r) are zero-order correlations between the predictor variables and reading times for each analysis region

* $p \leq .05$

Procedure

The participants were informed that they would be presented with a number of vignettes to read on a word-by-word basis. To advance through the vignettes, they pressed the “Next Word” button on a buttonbox. Dashes were used to represent the rest of each passage. Only one word was visible at a time. Comprehension questions appeared on 25 % of the trials. The vignettes were presented in a different random order for each participant. The participants completed two practice trials before beginning the actual experiment, which was run using E-Prime programming software, with a buttonbox to record participants’ reading times with millisecond accuracy.

Results

Comprehension accuracy was 94 %. Individual word reading times were analyzed at two points (see Example 1 below). The consequent wrap-up region (Region 1) was the earliest point at which the conditional could be evaluated as a whole. This was always the word or phrase immediately prior to the end of the consequent clause. We also measured reading times to the first word of the following sentence (Region 2) in order to reveal any residual (spill-over) processing load from Region 1 (Ehrlich & Rayner, 1983). Analysis on the reading time data was conducted using a linear mixed regression model with subjects and items as crossed random effects (see Locker, Hoffman, & Bovaird, 2007). The fixed effects included in the model as continuous predictors were the pretest ratings of $P(p)$, $P(q)$, $P(pq)$, $P(q|p)$, and $P(\text{not-}p \text{ or } q)$.

Example 1 The Union argues that if student tuition fees are increased, then applications for university places will / rise.¹ / At₂/ present university tuition fees can cost up to £3,000 per year.

Region 1 (wrap-up region)

The parameter estimates and p values (based on the t statistic) presented in Table 4 reveal that $P(p)$ and $P(q|p)$ significantly predicted reading times to this region. These variables were negative predictors, with decreased probability associated with increased reading time latencies (and vice versa). There was no association between pretest ratings of $P(q)$, $P(pq)$,¹ or $P(\text{not-}p \text{ or } q)$ and reading times to this region.

Region 2 (spill-over region)

The parameter estimates and p values (based on the t statistic) presented in Table 4 reveal no significant effects of probability on reading time for this region.

Discussion

In our reading time experiment, wrap-up latencies were predicted by pretest ratings of $P(p)$ and $P(q|p)$. There were no influences of $P(q)$, $P(pq)$, or $P(\text{not-}p \text{ or } q)$ on reading times for the critical region of the text. Our results suggest that readers use the subjective conditional probability to rapidly guide their interpretation of a conditional as it is comprehended in real time. This is consistent with predictions made by the suppositional theory (Evans & Over, 2004). No evidence emerged that during processing readers base their evaluation on an initial mental model representing the conjunction of p

¹ Because the product of $P(p)$ and $P(q|p)$ is logically equivalent to $P(pq)$ (Over et al., 2007), we performed a second analysis in which we entered $P(p)*P(q|p)$ into our model in place of $P(pq)$. As a validity check, we calculated the correlation between our pretest ratings of $P(pq)$ and $P(p)*P(q|p)$, finding that they were highly correlated ($r = .92$, $p \leq .001$). Consistent with our main analyses, this proxy for $P(pq)$ was not significantly associated with reading times to either of our analysis regions. The correlation between $P(p)*P(q|p)$ and $P(q|p)$ was .66, suggesting that collinearity was not an issue.

and q or on a fully fleshed-out model represented by the probability of the material conditional.

In addition to the effect of $P(q|p)$, we also found that $P(p)$ predicted wrap-up latencies. While this was not a primary prediction, it is nevertheless consistent with a version of the Ramsey test in which readers make a minimal change to their beliefs in order to hypothetically suppose the antecedent proposition (p) to be true (Stalnaker, 1968). In Stalnaker's version of the Ramsey test, beliefs about the actual world must be temporarily altered. For subjectively high-probability antecedents, this change in beliefs is likely to be negligible, but for low-probability antecedents, a much bigger and more cognitively demanding change in beliefs is required. Because $P(p)$ was a negative predictor, this effect most likely reflects the relative difficulty in updating the discourse representation to suppose improbable events as though they were true [with low $P(p)$ clauses associated with increased latencies].

Of the five sources of probability information that we manipulated, only $P(p)$ and $P(q|p)$ were associated with wrap-up reading times. For example, $P(q)$ did not predict latencies to the consequent (q) wrap-up region. In other words, readers processing the conditional *if student tuition fees are increased, then applications for university places will fall* were influenced by the probability of tuition fees increasing and the probability of applications falling given an increase in fees, but the probability of university applications falling in their own right did not matter. This suggests that not all sources of probability associated with conditionals are weighted equally in the mind of the reader.

Consistent with a number of previous studies, we found that $P(\text{not-}p \text{ or } q)$ had no influence on the evaluation of a conditional (Evans et al., 2003; Over et al., 2007). This again showed that the material conditional of propositional logic does not influence the psychological representation of conditional information. Unlike a subset of previous studies (Evans et al., 2003; Oberauer & Wilhelm, 2003; Politzer et al., 2010), we found no evidence that $P(pq)$ influenced the interpretation of our conditionals. This is unsurprising, given that the effect of $P(pq)$ has not been consistently replicated in the literature (Evans et al., 2008) and has only been shown in offline studies measuring the ultimate interpretation of a conditional. One speculative possibility is that the influence of $P(pq)$ found in previous studies may have been a remnant of sustained analytical processing. In contrast, the fast-acting heuristic processes required to rapidly evaluate a conditional online may be more suited to a mechanism whereby *if* immediately triggers a supposition, and the probability of q is evaluated only within this hypothetical world.

Our findings provide an initial insight into the processing of probabilistic information as readers rapidly establish their belief in indicative conditionals. Both the suppositional and

mental-models theories are well supported by evidence from offline tasks using abstract conditionals, but the offline techniques that have typically been employed within the reasoning literature have a limited capacity to advance our understanding of the processing of conditionals during incremental comprehension. Theoretical advances have recently been driven by a focus on the interpretation of everyday causal conditionals (e.g., Over et al., 2007). We believe that an examination of the online mechanisms associated with processing conditionals is also essential for refining existing theories and posing questions that have not previously been considered.

Author Note This research was supported by Leverhulme Trust Grant No. F/00 120/BT awarded to the second author.

References

- de Vega, M., Urrutia, M., & Rizzo, R. (2007). Cancelling updating in the comprehension of counterfactuals embedded in narratives. *Memory and Cognition*, *35*, 1410–1421.
- Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye-movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 75–87.
- Espino, O., Santamaria, C., & Byrne, R. M. J. (2009). People think about what is true for conditionals, not what is false: Only true possibilities prime the comprehension of “if. *Quarterly Journal of Experimental Psychology*, *62*, 1072–1078.
- Evans, J. S. B. T., Handley, S., Neilens, H., & Over, D. (2007). Thinking about conditionals: A study of individual differences. *Memory and Cognition*, *35*, 1772–1784.
- Evans, J. S. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 321–335.
- Evans, J. S. B. T., Neilens, H., & Over, D. (2008). Understanding causal conditionals: A study of individual differences. *Quarterly Journal of Experimental Psychology*, *61*, 1291–1297.
- Evans, J. S. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Ferguson, H. J. (2012). Eye movements reveal rapid concurrent access to factual and counterfactual interpretations of the world. *Quarterly Journal of Experimental Psychology*, *65*, 939–961.
- Ferguson, H. J., & Sanford, A. J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language*, *58*, 609–626.
- Fugard, J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 635–648.
- Giroto, V., & Johnson-Laird, P. N. (2004). The probability of conditionals. *Psychologica*, *47*, 207–225.
- Haigh, M., & Stewart, A. J. (2011). The influence of clause order, congruency and probability on the processing of conditionals. *Thinking and Reasoning*, *17*, 402–423.
- Haigh, M., Stewart, A. J., Wood, J. S., & Connell, L. (2011). Conditional advice and inducements: Are readers sensitive to implicit speech acts during comprehension? *Acta Psychologica*, *136*, 419–424.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove: Erlbaum.

- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.
- Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M., & Caverni, J.-P. (1999). Naïve probability: A mental model theory of extensional reasoning. *Psychological Review*, *106*, 62–88.
- Locker, L., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modelling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, *39*, 723–730.
- Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 680–693.
- Over, D. E., Hadjichristidis, C., Evans, J. S. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*, 62–97.
- Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking and Reasoning*, *16*, 172–197.
- Rader, A. W., & Sloutsky, V. M. (2002). Processing of logically valid and logically invalid conditional inferences in discourse comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 59–68.
- Ramsey, F. P. (1990). General propositions and causality. In D. H. M. Mellor (Ed.), *Philosophical papers* (pp. 145–163). Cambridge: Cambridge University Press (Original work published 1931).
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, *53A*, 1061–1080.
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1290–1301.
- Stalnaker, R. (1968). A theory of conditionals. *American Philosophical Quarterly Monograph Series*, *2*, 98–112.
- Stewart, A. J., Haigh, M., & Kidd, E. (2009). An investigation into the online processing of counterfactual and indicative conditionals. *Quarterly Journal of Experimental Psychology*, *62*, 2113–2125. doi:10.1080/17470210902973106