Logic as Marr's Computational Level: Four Case Studies

Giosuè Baggio,[a,b] Michiel van Lambalgen,[c] Peter Hagoort[d,e]

[a]*Brain and Language Laboratory, Neuroscience Area, SISSA International School for Advanced Studies*
[b]*Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology*
[c]*Department of Philosophy, Institute for Logic, Language and Computation, University of Amsterdam*
[d]*Donders Centre for Cognitive Neuroimaging,*
[e]*Max Planck Institute for Psycholinguistics*

Abstract

We sketch four applications of Marr's levels-of-analysis methodology to the relations between logic and experimental data in the cognitive neuroscience of language and reasoning. The first part of the paper illustrates the explanatory power of computational level theories based on logic. We show that a Bayesian treatment of the suppression task in reasoning with conditionals is ruled out by EEG data, supporting instead an analysis based on defeasible logic. Further, we describe how results from an EEG study on temporal prepositions can be reanalyzed using formal semantics, addressing a potential confound. The second part of the article demonstrates the predictive power of logical theories drawing on EEG data on processing progressive constructions and on behavioral data on conditional reasoning in people with autism. Logical theories can constrain processing hypotheses all the way down to neurophysiology, and conversely neuroscience data can guide the selection of alternative computational level models of cognition.

Keywords: Logic; Semantics; Cognitive neuroscience; Language; Reasoning; EEG

1. Introduction

Marr (1982) argued that cognitive systems should be understood at three independent levels of analysis: a *computational level*, specifying the goal and logic of the computation; an *algorithmic level*, providing a representation for the input and output, and an algorithm for the transformation; and a *level of physical implementation*, describing how the algorithms can be realized in physical machines. We present four case studies in the

---

cognitive science of language and reasoning, showing that computational level theories based on formal logic can predict and explain patterns of neurophysiological and behavioral data, and that such types of data can help with selecting among alternative computational theories.

2. The explanatory power of logic in neuroscience

2.1. Bayesian and logical theories of conditional reasoning

To show how neural data can guide the selection between computational level theories, we compare probabilistic and logical theories of conditional reasoning. The Bayesian treatment considers a conditional to be semantically interpretable as a conditional probability. Logical models favor some kind of defeasible (i.e., non-monotonic) logic in which certain conclusions may be withdrawn if a disabling condition is added to the initial premises, as in the example below. Both models aim to be computational level theories of behavior on tasks such as *modus ponens* (MP) reasoning. In Pijnacker et al. (2011), there are two conditions, neutral and disabling:

MP-Neu

Lisa has recently bought contact lenses (neutral condition *N*)

If Lisa is going to play hockey, she will wear contact lenses (conditional premise)

Lisa is going to play hockey (categorical premise)

Lisa will wear contact lenses (conclusion)

MP-Dis

Lisa probably lost a contact lens (disabling condition *D*)

If Lisa is going to play hockey, she will wear contact lenses (conditional premise)

Lisa is going to play hockey (categorical premise)

Lisa will wear contact lenses (conclusion)

In the study by Pijnacker, Geurts, van Lambalgen, Buitelaar, and Hagoort (2011), participants endorsed *modus ponens* with a neutral condition in $\sim 90\%$ of the cases, as opposed to $\sim 50\%$ with a disabling condition. This is one variant of the so-called suppression effect (Byrne, 1989). In the original task, the possibly disabling information comes in the form of a conditional second premise, which would in this case be: "If Lisa hasn't lost her contact lenses, she will wear them."

The original "suppression task" was used by Oaksford and Chater (2003) to argue for a Bayesian model of conditional reasoning and against a logical approach. Bayesian probability is said to be the proper computational level theory for conditional reasoning because conditionals are never certain and "probability is the calculus of uncertainty."[1] Moreover, it is suggested that the Bayesian treatment fits the data more closely than any logical analysis.

Below we provide Marr-type analyses of logical and Bayesian reasoning in this conditional task and compare these to ERP (event-related potential) data on the time course of processing.

2.1.1. Logical analysis at the computational level

Our aim here is to represent the premises in such a way that inferences can be drawn (or not drawn). The appropriate notion of validity for non-monotonic reasoning is as follows: view the premises of the argument as a discourse; this needs to be integrated, and a coherent discourse model must be constructed; it then has to be computed whether the putative conclusion holds in the discourse model.

2.1.2. Logical analysis at the algorithmic level

Here processing considerations come to the fore, in particular the "principle of immediacy"—the notion that computation "starts immediately with what it derives on the basis of the bottom-up input and the left context" (Hagoort, 2007).

For the disabling condition the computation proceeds as follows, where $\neg ab$ means nothing abnormal is the case:

1. D [disabling condition: "Lisa probably lost a contact lens"];
2. H [categorical premise: "Lisa is going to play hockey"];
3. $H \wedge \neg ab \rightarrow W$ [conditional; W is "Lisa will wear contact lenses"];
   $D \wedge \neg ab' \rightarrow ab$ [integrating disabling condition and conditional];
4. $?W$ [is W true in discourse model?]; it is argued in van Lambalgen and Hamm (2005) that computation here proceeds "backwards":
5. W is true if $H \wedge \neg ab$ is true; we know H is true, and if D constitutes an unconditional abnormality for $H \wedge \neg ab \rightarrow W$ (meaning that $\neg ab'$ holds), ab follows, and the antecedent $H \wedge \neg ab$ cannot be satisfied, whence W cannot be made true in the discourse model. On the other hand, if it is judged that D constitutes an abnormality only in special circumstances, we have that the antecedent of $D \wedge \neg ab' \rightarrow ab$ is false, hence so is the consequent ab, from which it follows that W;
6. The computation in the neutral condition N does not introduce $D \wedge \neg ab' \rightarrow ab$ together with D; since we don't have information about ab, we assume $\neg ab$, and the *modus ponens* argument goes through.

The supposition $\neg ab$ is an instance of the closed world assumption, which allows one to take as true the negation of all propositions which are not entailed by the premises and by background knowledge.

2.1.3. Bayesian analysis at the computational level

As in the logical theory, we see defeasible reasoning as involving the construction of discourse models. In the Bayesian account, however, discourse models carry a joint probability distribution specifying, for instance, the conditional probability that Lisa will wear lenses given that she is going to play hockey. The modus ponens inference is represented by Bayesian conditionalisation:

if B is a complete (non-probabilistic) description of one's knowledge, the posterior probability $P_1(A)$ given that B has occurred is equal to the prior conditional probability $P_0(A|B)$,

2.1.4. Bayesian analysis at the algorithmic level

As a constraint we have again the principle of immediacy. The computation then unfolds as follows:

1. D [disabling condition], this introduces the conditional probability $P(\bullet|D) =: P(\ )_D$;
2. The implication "if H then W" is represented by a conditional probability with respect to the distribution $P(\ )_D$, as follows:

$$\text{if H then W} \iff \frac{P(H \wedge W)_D}{P(H)_D} = P(W|H)_D = P(W|H, D).$$

Immediacy dictates that $P(\ )_D$ (which gives the *posterior* conditional probability) is used instead of P (the *prior* conditional probability);

3. H;
4. In the last step one conditionalises on H in the expression $P(W|H)_D$, and one decides on one of the three answers "yes, maybe, no";
5. The computation for N [neutral condition] is in principle the same, except that one may assume $P = P(\ )_N$.

A more refined analysis of the disabling condition would not conditionalise on D, since this event is only said to be probable; instead, one applies *Jeffrey conditionalization* to obtain as posterior conditional probability:

$$\text{if H then W} \iff \frac{P(H \wedge W)_D}{P(H)_D} P(D) = \frac{P(H \wedge W)_{\neg D}}{P(H)_{\neg D}} P(\neg D).$$

2.1.5. Comparing the logical and the Bayesian analyses

On the logical model, the evaluation of the conclusion is both computationally intensive and the source of competing representations. In the Bayesian analysis, the computation of the posterior conditional probability in the disabling case requires most resources,

whereas settling upon an answer from "yes, maybe, no" requires the subject to set a criterion, but no further computation is necessary. This yields the following predictions for ERPs:

(i) on the Bayesian analysis, one expects an ERP effect in the disabling case as compared to the neutral case when the *conditional premise* is processed; as neural learning algorithms, Bayesian models aim at minimizing prediction errors (Friston & Stephan, 2007); violations of default processes may lead to increasing prediction errors reflected by ERP effects such as a mismatch negativity (Garrido, Kilner, Kiebel, & Friston, 2007; Kiebel, von Kriegstein, Daunizeau, & Friston, 2009) or a P300;

(ii) on the logical model, the ERP effect should occur while the *conclusion* is processed; the most likely outcome here is the kind of sustained negativity observed for instances of defeasible inference and referential ambiguity in discourse processes (Baggio, van Lambalgen, & Hagoort, 2008; van Berkum, Zwitserlood, Hagoort, & Brown, 2003).

One also expects differences in the behavioural domain. The task is sufficiently analogous to the original suppression task to expect a decrease in MP endorsements of about 40%, perhaps a bit more because the disabling condition is introduced categorically, not hypothetically. It is not so clear what to expect on the Bayesian model, but the following line of reasoning seems plausible: Since the categorical disabling condition is said to be probable, the posterior conditional probability of "if H then W" must be low, hence there should be a high percentage of "maybe" answers.

Pijnacker et al. (2011) designed an ERP study to establish the time-course of brain processes subserving defeasible reasoning. Signals were recorded while participants read a set of four sentences (of type MP-Dis or MP-Neu), and pressed keys corresponding to "yes," "no" or "maybe" after reading the conclusion sentence.

ERP averages were computed over all sets of type MP-Dis, and likewise over all sets of type MP-Neu. The average MP-Neu was then compared to the average MP-Dis at different positions (see below) within the discourse. The results were as follows:

1. no significant difference between average MP-Neu and average MP-Dis for the first sentence (neutral or disabling condition);
2. no significant difference between average MP-Neu and average MP-Dis for the conditional premise;
3. no significant difference between average MP-Neu and average MP-Dis for the categorical premise;
4. a significant long-lasting negative ERP shift of average MP-Dis compared to average MP-Neu for the conclusion, starting 250 ms after the onset of the final word of the conclusion.

These results show that, until the onset of the final word of the conclusion, the computational load does not differ between neutral and disabling cases. From that point onward, processing costs for the disabling case are higher than for the neutral case. These results conflict with the probabilistic model outlined above. The Bayesian model predicts a heavy processing load after the second premise, and no processing

load in the final step, which is just Bayesian conditionalisation. The logical model looks a bit better, and it may even provide a reason why a sustained anterior negativity is observed. We noted that the appearance of the conclusion triggers competition between representations, and it seems of some interest to observe that the sustained negativity reported by Pijnacker et al. (2011) is similar to the one observed with ambiguous referents, as in "David had told the two girls to clean up their room before lunchtime. But one of girls had stayed in bed all morning, and the other had been on the phone all the time. David told *the girl* that..." (van Berkum et al., 2003)—here the italicized NP is referentially ambiguous, leading to competition between the representations of each referent.

2.2. Processing temporal connectives: A logical model

We have just seen how experimental data can be used to select between alternative accounts (Bayesian probability or defeasible logic) of conditional reasoning. We will now show that computational level analyses inspired by logic can provide alternative explanations of existing data, identify potential confounds, and guide predictions for follow-up experiments. An example of this is a study by Münte, Schiltz, and Kutas (1998) in which ERPs were recorded while subjects silently read sentences that differed only in the temporal connective, for example:

(A) After the scientist submitted the paper, the journal changed its policy.
(B) Before the scientist submitted the paper, the journal changed its policy.

"Before" sentences elicited a larger sustained negativity. The amplitude was largest over left anterior electrodes, where neural responses to "before" and "after" diverged at 300 ms from word onset. The negative shift was largest during the second clause.

According to Münte et al., the observed ERP patterns can be accounted for as follows. When in sentence-initial position, "after" signals that events are presented in chronological order, whereas "before" indicates that event order is reversed. It is suggested that "before" sentences require additional computations as a consequence of presenting the events in reverse order. However, a computational level theory of temporal connectives must also consider the asymmetry between the entailment properties of "before" and "after." Whereas (B) can be taken to mean that the scientist was not able to submit her paper because of the journal's policy change, this meaning is not available with "after." This flexibility is referred to as the "non-veridicality" of "before" and is a potential confound in this study (Baggio, 2004).

2.2.1. Computational level
We treat the asymmetry of "after" and "before" as involving the anchoring of events on a timeline. The main event is always bound to a time interval. "After" demands that the subordinate clause event is also bound to the timeline, whereas "before" allows for this event to be left unanchored or "floating": no commitment is made as to whether, and when, that event happens. This informal definition can now be applied to Münte et al.'s

stimulus sentences. In (A), the scientist submitted the paper at time t, and the journal changed its policy at t' later than t. In (B), in contrast, the journal's policy change at time t may prevent the scientist from submitting her paper at any later t', for example, if the article is judged to be unsuitable given the journal's new policy (i.e., a disabling condition). What is required by the semantics of "before" is that the submission event does not happen at any time $t' < t$. Therefore, t' is left floating on the timeline. This analysis can be fully expounded in the logical calculus of events by van Lambalgen and Hamm (2005). Here we provide a somewhat less formal description of the interpretation of these sentences at the algorithmic level.

2.2.2. Algorithmic level

Let e be the event "the scientist's submission of a paper," and e' the event "the journal's change of policy." The interpretation of "Before e, e'" is given by the following algorithm:

1. put the event e in a temporary store;
2. update discourse model with e' situated in the past, and mesh with already present events;
3. compute states consequent upon e';
4. check whether any of these consequent states conflicts with preconditions for e;
5. if so, empty the store, but do not update discourse model with e;
6. if not, empty the store, update discourse model with e, situated between e' and "now."

By contrast, the meaning for "After X, Y" is always veridical, and therefore the events e, e' (corresponding to the clauses X and Y) can be immediately integrated in the discourse model:

1. update the discourse model with event e and time t in the past and mesh e with already present events;
2. update the resulting discourse model with event e' and time s later than t (how much later depends on the particular events).

In "Before e, e'" constructions only, to update the current model based on information provided by the subordinate clause event e, the system has to hold on for the main clause event e' to be interpreted, regardless of whether, while processing the "before" clause, it entertains tentative interpretations or not. If the execution of these updates is indeed delayed, the event e has to be held in working memory until relatively late (partly depending on the position of the main clause verb) before it can be evaluated. This might explain the ERP effects reported by Münte et al. (1998): Could it be that the interpretation of "before" clauses, rather than event order, is what increases processing load? Logic can inspire alternative computational level theories which may carry along alternative explanations of processing data, as well as ideas for removing the confounds thus identified.

3. The predictive power of logic in neuroscience

3.1. The progressive: Processing predictions from a logical model

So far we have considered two cases in which computational level theories come on the scene after the data have been collected, and thus have no influence on shaping predictions and experimental designs—as though logic played the role of Minerva's owl in cognitive neuroscience. In what follows, we show that computational level theories based on logic can guide experimentation by providing explicit processing predictions. Consider the sentence

(*N*) The girl was writing a letter when her friend spilled coffee on the tablecloth.

From (*N*) the reader would conclude that, barring unforeseen obstacles, the girl will attain the desired goal (a complete letter) and would assent to the statement "The girl has written a letter" (Baggio & van Lambalgen, 2007). This inference is based on knowledge that spilling coffee on the tablecloth is not a typical (disabling) event that leads to the termination of the writing process. In fact, the closed world assumption allows one to conclude there are no disabling events in the scenario, as none follows from the discourse. Finally, a principle of inertia ensures that the writing activity will reach its natural endpoint in finite time.

The inference to the goal state is non-monotonic—it can be suppressed if discourse introduces an event which does terminate the relevant activity:

(*D*) The girl was writing a letter when her friend spilled coffee on the paper.

Assuming that writing was intended to occur on the same paper sheets on which coffee was spilled, the accident is a disabling condition to obtaining a complete letter. Accordingly, upon reading (*D*) the reader would assent to "The girl has written no letter." As we found in two subsequent behavioral studies (Baggio & van Lambalgen, 2007; Baggio et al., 2008), endorsement rates of goal state conclusions ("The girl has written a letter") were higher in the neutral condition (*N*) (74.925% and 76.05% in the 2007 and 2008 studies, respectively) compared to the disabling condition (*D*) (37.15% and 54.68%). This entailment pattern resembles the "suppression effect" in reasoning with conditionals discussed above, and indeed the underlying (non-monotonic) computational level logic is the same.

At the algorithmic level, the semantics of "x was writing a letter, when ϕ" is given by the following series of instructions:

1. update the discourse model with a temporal interval I corresponding to the extended activity "writing," where I contains a past reference time R established by the given discourse model;
2. update the discourse model with a letter L and a time $t > R$ at which L is expected to be finished (L,t);

3. update the discourse model with the event *e* corresponding to the subordinate clause *φ* (we may assume *e* occurs at *R*);
4. compute states consequent upon *e* which bear upon the existence of *L*;
5. if one of these states conflicts with the existence of *L* at *t*, remove *L,t* and recompute the discourse model.

This set of instructions can be formalized in constraint logic programming in ways analogous to the treatment of conditional reasoning sketched above.

Condition (*D*) satisfies (5) and so is expected to elicit a recomputation of the discourse model when reading the disabling clause. Baggio et al. (2008) conducted an ERP study in which it was found that the sentence-final word in (*D*) indeed produced a larger sustained anterior negativity relative to the sentence-final word in (*N*). The amplitude of this ERP effect was inversely correlated with the rate of goal state conclusions ("The girl wrote a letter") as assessed by button-press responses after each trial.

This result suggests that computational level theories grounded in logic can provide processing predictions that can be tested with psycholinguistic methods such as ERPs. We emphasize that what can be asked to a cognitive theory is to generate definite processing predictions (theory-to-data route), whereas ruling out alternative explanations is the concern of experimental research through control and follow-up studies (data-to-theory route). It is, however, sometimes possible to examine alternative explanations of a given processing phenomenon by means of theoretical considerations at the level at which the explanation is said to hold. For example, in Baggio et al. (2008) it is argued that an "alternative" monotonic account of the progressive based on possible worlds semantics is in fact indistinguishable from the present analysis vis-á-vis ERP data, as it too predicts a recomputation of the initial (lower) probability states, associated with the possible worlds in which the goal state is not attained, when the disabling clause is read. It therefore falls to both theory and experimentation to jointly develop "crucial experiments" in which alternative theories can effectively be disentangled, besides single studies designed to test the predictions of a single theory.

3.2. Logic and executive (dys)function in autism

The fourth case study is concerned with the use of logic as a computational level model of executive *dys*function in cognitive disorders. Executive function comprises planning, initiation, inhibition, coordination, and control of action sequences leading to a goal retained in working memory. Here we abstract from the control and co-ordination components, and focus on goal maintenance, planning and (contextually determined) inhibition. By definition, planning consists of the construction of a *sequence* of actions which will achieve a given goal, taking into account properties of the world and the agent, and also events that might occur in the world. A major problem with planning is that it is in general impossible to take into account all eventualities whose occurrence might be relevant to the success or failure of the plan. Hence, it is crucial to be able to plan flexibly so that the execution of a sequence of actions can be inhibited if changed circumstances demand this.

In the logical model of executive function proposed here, inhibition is represented through the special logical form of the conditional we saw in section 2.1, where the link from condition (C) to action (A) is mediated by a slot labeled $\neg ab$:

$$C \wedge \neg ab \rightarrow A$$

This is read as "if C and *nothing abnormal is the case*, then A." The expression "*nothing abnormal is the case*" is governed by closed world reasoning, and the abnormality acts as an inhibitor: If ab holds, the connection between C and A is interrupted.

Russell's (1997) executive disorder theory takes as basic the observation that autistic people often exhibit perseveration, that is, they go on carrying out a routine when that routine is no longer appropriate, and they show great difficulties in switching tasks when the context calls for this, that is, when switching is not governed by some explicit rule. Russell drew attention to the similarity between flexible planning and defeasible reasoning:

> [T]aking what one might call a "defeasibility stance" towards rules is an innate human endowment—and thus one that might be innately lacking ... [H]umans appear to possess a capacity—whatever that is—for abandoning one relatively entrenched rule for some novel ad hoc procedure. The claim can be made, therefore, that this capacity is lacking in autism, and it is this that gives rise to failures on "frontal" tasks—not to mention the behavioral rigidity that individuals with the disorder show outside the laboratory. (Russell, 2002, p. 318)

This suggested to us that one might test people with autism on a defeasible reasoning task such as the one discussed in section 2.1. Our prediction here was that there should be a difference between people with autism and controls in the condition MP-Dis, in that the former group should be far less inclined to suppress the *modus ponens* inference. This was indeed observed. In both groups the baseline MP-Neu scored around 90%, but MP-Dis in controls scored 51%, whereas in the autism group 71% ($p = 0.025$).[2] Here, a logical theory of executive function at the computational level, together with data on executive dysfunction in autism, led to a novel prediction on reasoning with exceptions in people with autism. This was possible because the logic underlying the computational level theory for reasoning with exceptions is very similar to that of the computational level theory for flexible planning.

4. Conclusion

Marr (1982, p. 25) defined the computational level as providing answers to questions such as:

What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?

Computational theories include a description of the "why" of cognition ("the goal of the computation") (Willems, 2011), as well as *a specification language for the algorithms* ("the logic of the strategy by which it can be carried out"), and that is what formal logic can contribute to the cognitive neuroscience of language and reasoning. In these domains, compared to other formalisms such as Bayesian probability, logical theories exhibit sufficient explanatory and predictive power to adequately model aspects of behavioral and neural data. It is beyond the scope of this paper to argue for the supremacy of logic as a computational level theory in cognitive domains other than discourse processing and reasoning—but see beim Graben and Potthast (2009) for a broader stance than ours. It seems reasonable to predict that cognitive tasks involving inference will afford some logical analysis which may potentially be supported by data. Candidate areas include language production and language acquisition in so far as they require representations of events and other complex semantic objects. Conversely, there are phenomena that may not prove fertile ground for logics, or for computational level analyses tout-court but may be entirely explained by models at the level of physical implementation—for example, free conceptual associations (Russo, Namboodiri, Treves, & Kropff, 2008) and automatic semantic priming (Lerner, Bentin, & Shriki, 2012). Ultimately, however, it is the task of experimental research to trace the boundaries of the information processing tasks for which computational level analyses can be carried out, and within those boundaries, what the province of logical analyses may be.

Notes

1. For a detailed discussion of the hypotheses that can be used to justify this assertion, see Paris (1994).
2. For details concerning the experiment, we refer to Pijnacker (2009).

References

Baggio, G. (2004). Two ERP studies on Dutch temporal semantics. MoL Series, Institute for Logic, Language and Computation.

Baggio, G., & van Lambalgen, M. (2007). The processing consequences of the imperfective paradox. *Journal of Semantics*, *24*, 307–330.

Baggio, G., van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, *59*, 36–53.

beim Graben, P., & Potthast, R. (2009). nverse problems in dynamic cognitive modeling. *Chaos*, *19*, 015103.

Byrne, R. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 61–83.

Friston, K., & Stephan, K. (2007). Free-energy and the brain. *Synthese*, *159*, 417–458.

Garrido, M., Kilner, J., Kiebel, S., & K. Friston (2007). Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences 104*, 20961–20966.

Hagoort, P. (2007). The memory, unification, and control (MUC) model of language. In A. S. Meyer, L. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 243–270). Hove: Psychology Press.

Kiebel, S., von Kriegstein, K., Daunizeau, J., & Friston, K. (2009). Recognizing sequences of sequences. *PLoS Computational Biology*, *5*, e1000464.

Lerner, I., Bentin, S., & Shriki, O. (2012). Spreading activation in an attractor network with latching dynamics: Automatic semantic priming revisited. *Cognitive Science*, *36*, 1339–1382.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Múnte, T., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, *395*, 71–73.

Oaksford, M., & Chater, N. (2003). Probabilities and pragmatics in conditional inference: Suppression and order effects. In D. Hardman and L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision* (pp. 95–122). London: Wiley

Paris, J. (1994). *The uncertain reasoner's companion*. Cambridge, UK: Cambridge University Press.

Pijnacker, J. (2009). Defeasible reasoning in high-functioning adults with autism: Evidence for impaired exception-handling. *Neuropsychologia*, *47*, 644–651.

Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J. & Hagoort, P. (2011). Reasoning with exceptions: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, *23*, 471–480.

Russell, J. (1997). *Autism as an executive disorder*. Oxford, UK: Oxford University Press.

Russell, J. (2002). Cognitive theories of autism. In J. Harrison & A. Owen (Eds.), *Cognitive deficits in brain disorders*, (pp. 295–323). London: Dunitz.

Russo, E., Namboodiri, V., Treves, A. & Kropff, E. (2008). Free association transitions in models of cortical latching dynamics. *New Journal of Physics*, *10*, 015008.

van Berkum, J., Zwitserlood, P., Hagoort, P. & Brown, C. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*, 701–718.

van Lambalgen, M., & Hamm, F. (2005). *The proper treatment of events*. Oxford, Blackwell.

Willems, R. (2011). Re-appreciating the why of cognition: 35 years after Marr and Poggio. *Frontiers in Psychology 2*(244), 1–5.