Probabilities, Causation, and Logic Programming in Conditional Reasoning: Reply to Stenning and Van Lambalgen

Mike Oaksford*

Birkbeck College, University of London

Nick Chater

Warwick Business School, University of Warwick

Short Title: Probabilities, Causation, and Logic Programming in Conditional Reasoning

*Corresponding author:

Mike Oaksford

Department of Psychological Sciences

Birkbeck College

University of London

London, WC1E 7HX, UK

e-mail: Mike.Oaksford@bbk.ac.uk

telephone: +44 (0) 20 7079 0879

Abstract

Oaksford and Chater (2014) critiqued the logic programming (LP) approach to nonmonotonicity and proposed that a Bayesian probabilistic approach to conditional reasoning provided a more empirically adequate theory. The current paper is a reply to Stenning and van Lambalgen's (in press) rejoinder to this earlier paper. It is argued that causation is basic in human cognition and that explaining how abnormality lists are created in LP requires causal models. Each specific rejoinder to Oaksford and Chater's (2014) original critique is then addressed. While many areas of agreement are identified, with respect to the key differences, it is concluded the current evidence favours the Bayesian approach, at least for the moment.

In their rejoinder to our paper in the recent special issue on dual processes in this journal, Stenning and Van Lambalgen (2015, henceforth "SL") argue that we have misrepresented their logic programming approach (henceforth, "LP") to human reasoning. We should say at the outset that we are grateful to SL for the constructive nature of their response and that we hope that our reply will be interpreted in a similarly constructive vein. Following SL, we focus our response on our argument that causal Bayes nets (CBNs) may provide a useful formalism for theorising about conditional inference (Ali, Chater, & Oaksford, 2011; Ali, Schlottmann, Shaw, Chater, & Oaksford, 2010; Fernbach & Erb, 2013; Chater & Oaksford, 2006; Oaksford & Chater, 2010a, 2013, 2014, in press). However, the application of probability theory to the psychology of human reasoning goes beyond this recent proposal. There is currently a range of probabilistic approaches, which go under the rubric of *The New Paradigm* (Manktelow, 2012; Over, 2009), which have many commonalities but some important differences (e.g., Baratgin, Over, & Politzer, 2013; Douven & Verbrugge, 2013; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Gilio & Over, 2012; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007; Politzer, 2005; Pfeiffer & Kleiter, 2010). In particular, our approach has been Bayesian from the outset (Oaksford & Chater 1994) but has also been influenced by Adams (1998) work on probability logic and work on CBNs (Pearl, 1988; 2000). In this reply, while focusing on CBNs, we will appeal to a range of probabilistic approaches taken in the literature. This is because it is clear from SL's response that the use of probabilities in theorising about human reasoning and language interpretation is the primary target of their rejoinder.

Some Intuitions

We first outline two intuitions that guide our approach to the psychology of conditional reasoning. We then respond to the points that SL raise.

It's causal, all the way down.

The origins of the main contemporary accounts of the conditional are based on intuitions about causal or law-like dependencies. In the philosophy of science, it was soon realised that the material conditional ($p \supset q$), which is true as long as p is false or q is true, could not account for such dependencies (Chisolm, 1946; Goodman, 1947, 1955). Law-like relations *support counterfactuals*. That is, the reason we believe that the counterfactual (1),

If the key had been turned the car would have started (1)

to be true is that we believe that the indicative (2)

If the key is turned the car starts (2)

describes a real causal dependency in the world (Stalnaker, 1984). Notice that according to the material conditional, counterfactuals are always true, because their antecedents, i.e., p, are always false. This is, after all, why they are *counter*factuals. Consequently, the material conditional could not capture our intuitions about counterfactuals.

The possible worlds semantics for the counterfactual (Lewis, 1973; Stalnaker, 1968) is widely recognised as the beginning of our understanding of the conditional of natural language as opposed to in mathematics (Nute, 1984). Agreeing with SL on the importance of the Ramsey test, both probability logic (Adams, 1998) and the possible worlds semantics for conditionals (Stalnaker, 1968, 1984) have their origins in this test. For example, the truth of (1) involves considering the possible world most similar to the real world in which the key was turned, if the car starts in this world then the conditional is true. Exactly the same formulation applies to "open" indicative conditionals where it is not known whether the antecedent is true or false (2). Stalnaker's (1984) conceptualist interpretation of possible worlds also indicates how causality is fundamental to our evaluation of conditionals. On this interpretation, open indicative conditionals correspond to our methodological policies for changing our beliefs. As we have seen, if someone asserts (1) they must believe that the methodological policy in (2) describes a real causal dependency in the world. This is a claim about the person's psychology: we can remain neutral on the ontological question of whether such a relation really exists. All that matters is that our methodological policies, or as Hume referred to them our habits of inference, are believed to denote relations in the world that behave like causal relations. This is why we have argued that these habits of inference and hence the conditionals used to describe them can be represented in the structure of causal Bayes nets (Ali, Chater, & Oaksford, 2011; Chater & Oaksford, 2006; Oaksford & Chater, 2013). From this point of view, the way people think about the world is causal "all the way down" and causality is not reducible to anything else, like probabilities or similarity between possible worlds (Oaksford & Chater, 2010a, in press) .[1]

Some caveats are in order. Causal Bayes nets and the probabilistic inferences they allow are useful because there is uncertainty about the nature and efficacy of the causal relations that are operative in any given situation. The structure of a network relies on our causal intuitions but it and the strength of causal relations can be learned from data such as the direct observation of occurrences of putative causes and effects (Griffiths & Tenebaum, 2005), although this relies on some strong assumptions (i.e., faithfulness). But the input people receive may also consist of testimony from other people in which a causal or law-like relations are described using conditionals.

The received view of dual processes.

While there may be some room for disagreement, given the largely informal discussions of dual processes, there is a generally received view of System 1 as associative and heuristic and System 2 as being rule bound and tied to language. One might initially question how

[1] As Cartwright (1983) observes, any probabilistic analysis ends up having to define a probabilistic relation that keeps all other causally relevant factors constant, i.e., it is circular as a definition. As Stalnaker (1984) observes degrees of similarity and difference between possible worlds depends on similarities with respect to causal properties and so again circularity beckons.

associative and heuristic go together but the availability heuristic provides a clear example. This heuristic emerges as the result of world knowledge in LTM being represented in associative or connectionist networks in which the speed of retrieval is related to the strength of connections. Speed or ease of retrieval then provides a heuristic for judged probability. Given the body of evidence cited to support the basis of this division between these two systems (see, Evans, 2009; Sloman, 1995; Stanovich, 2011), a broad division of this kind seems a reasonable starting point for discussion, although many of the details may be contested. Indeed, a recent paper by Sternberg and McClelland (2012) provides relevant evidence for this type of division in a context closely related to the present analysis. Using a human contingency learning paradigm, they found that, under speeded conditions, their results were consistent with a pathway strengthening process like that implemented in connectionist or associative networks. However, with no time constraints, performance was consistent with an inferential approach based on choosing between possible causal models.

Response to Stenning and Van Lambalgen

As we just pointed out, the fundamental intuition about our use of CBNs to represent conditional statements was the relation to a non-reductive sense of causality. In this regard, LP and our approach are formally similar. This is because LP has been used to implement the Event Calculus (Kowalski & Sergot, 1986) and under certain conditions, “the event calculus would correspond to directed acyclic graphs that are equivalent to Pearl's causal structures” (Sowa, 2000/2015). In the event calculus, causation is also treated non-reductively. Bayesian networks and the event calculus provide alternative representations of causal relations, in just the same way that syntactic mental logics and mental models provide alternative mental representations of logical relations. Our contention is that the empirical evidence supports the psychological reality of CBNs and that there is a lack of empirical support for the LP approach. This is borne out in the first section of SL's rejoinder, which, after a brief explication of the LP approach, looks at applications of LP in the empirical literature.

Applications of LP.

The list of applications of LP to empirical phenomena is not long. This contrasts with the Bayesian cognitive science framework, which we and others apply to human reasoning. The empirical reach of this framework is evidenced by edited collections spanning the last 20 years (e.g., Chater & Oaksford, 2008; Oaksford & Chater, 1998).

In this section, SL also provide an example of how LP accounts for suppression effects. They then describe some empirical findings (Pijnacker et al, 2009, 2010) that replicate the suppression effect for the MP inference. We first critique SL's example as it is provides a confusing illustration of how LP works which needs clarification. The example inference is as follows (SL, p. ??):

> "MP-Disabling
> Lisa probably lost a contact lens. L (disabling condition)
> If Lisa is going to play hockey, she will wear contact lenses. H ∧ ¬ab → (conditional)
> L →ab (hidden premise from abnormality list expressing that L is disabling)
> Lisa is going to play hockey. H (categorical premise)
> Lisa will wear contact lenses. W (conclusion)
>
> MP-Neutral
> The formalisation is exactly the same except that there is no hidden premise L →ab, because buying contact lenses is not on the ab list."

This example is confusing because the disabling condition and the consequent of the conditional/conclusion (*W*) share content. SL describe *L* (probably losing a contact lens) as disabling for playing hockey (*H*). But regardless of whether Lisa plays hockey, *L* is not consistent with the conclusion *W*, that she will wear contact lenses. *Probably* losing a contact lens is as disabling for wearing contact lenses as it is for playing hockey. Indeed, people may regard having lost one's contacts as inconsistent with wearing them (conceivably, perhaps, one might hunt for one's contact lenses, without realizing one is wearing them, but the possibility seems remote). Consequently, the conclusion seems to be contradicted without the need to contemplate whether Lisa is playing hockey or not. That is, non-monotonic reasoning need have nothing to do with this inference.

Moreover, SL describe the *MP-Neutral* condition as simply missing the *hidden* premise $L \rightarrow ab$. By being "hidden" we assume this premise was not mentioned in the task and that it is just part of the computation in *LP*. But this raises the question: what is the difference in the materials presented to participants between the *MP-Neutral*[2] and the *MP-Disabling* conditions? Normally, using the explicit Byrne (1989) paradigm, *L* would not be presented in the *MP-Neutral* condition but the text seems to suggest that it was—apparently only the hidden premise $L \rightarrow ab$ was omitted in the computation. If *L* was presented in the *MP-Neutral* condition, then psychologically it seems odd that people could not infer from world knowledge that *L* is disabling for Lisa playing hockey, whether it was on an abnormality list for the conditional rule or not. Moreover, if it was not, then why is explicit mention of *L* required to put it on the abnormality list? For this conditional, where the consequent *W* shares content with the disabling condition, considering the possibility that she loses her contact lens yields the conclusion, she will not play hockey, by MT. That is, it is clear that losing a contact lens is disabling for Lisa playing hockey. The problems with this example show the care that needs to be taken in setting up genuine instances of defeasible reasoning. We suggest that the original example we borrowed from Wernicke better exemplifies how LP handles suppression effects.[3]

LP and Dual Process Theories.

[2] We do not understand the purpose of the reference to "*buying* contact lenses" in the description of the *MP-neutral* condition.

[3] Although this is an actual example of the materials used in Pijnacker et al. (2010), which may lead to questions over their results. However, unlike this example, Pijnacker et al. (2010) used a neutral context sentence in place of *L*.

In this section, SL first express their bafflement at our proposal that Stenning and van Lambalgen (2005) gives the impression that LP holds out the promise that non-monotonic reasoning can be handled by purely local System 2 computations. We are equally puzzled: the example of how LP explains the suppression effect in our original paper (Oaksford & Chater, 2014, p. 275, programs A and B) was taken verbatim from Stenning and van Lambalgen (2005, pp. 942-943). These programs require only a few local steps in order to draw suppression inferences. However, SL suggest that only "a cursory reading" of the literature on LP is required to encounter the publication SL cite (i.e., van Lambalgen & Hamm, 2004), [4] on how knowledge importation, i.e., System 1, works in LP. However, Stenning and van Lambalgen (2005) remains the primary source for the relevance of LP to psychological phenomenon such as suppression effects and we stand by our comments as an important corrective to the mistaken impression to which the discussion in the 2005 paper can easily give rise. Moreover, in our original paper we also stressed that LP can handle the computational complexities of knowledge importation. What we questioned was the psychological reality of the LP approach. This is the criticism to which SL turn next.

SL quote us saying, "…, how plausible is it to propose that people are explicitly indexing exceptions as abnormal." Their counterargument is to suggest that what Cummins (1995) refers to as "disablers" are just abnormalities and thus the effects observed are therefore, by definition, evidence for the psychological reality of abnormalities. But this is too quick. Our argument was against the psychological reality of abnormality propositions or predicates, *ab*, indexing a proposition or a subject as abnormal with respect to a rule. In a realistic data base, there have to be large numbers of different *ab* propositions/predicates, one for each rule. These abnormality propositions/predicates are a part of the deep logical structure underlying the interpretations of almost all conditionals. These are the LP

[4] Van Lambalgen and Hamm (2004) is a book implementing the event calculus in LP. It is a logico-linguistic work, which few psychologists of reasoning behaviour are likely to have read, bar S and L.

representations of our conditional beliefs. Yet, as we argued in our original paper, these abnormality propositions/predicates leave no trace in the surface linguistic forms or in our natural way of speaking about these beliefs.[5]

As we indicated in the opening section, we think causality comes first. We can see why by considering the role of the abnormality list and how items get onto this list. According to SL, the *MP-Neutral* case just needs to miss out the $L \rightarrow ab$ premise for MP to go through. Let us assume that for a particular reasoner L is not on their abnormality list, that is, the $L \rightarrow ab$ premise is missing. If L is then presented along with the conditional and categorical premises, how does L get on to the participant's abnormality list, as common sense seems to indicate it should? The way to do this is to build a causal model of the situation and simulate whether the presence of L prevents Lisa wearing contact lenses. In SL's example, where L and W share contents—they are both about contacts lenses—this is straightforward but it is less so for more realistic examples.

The need to construct causal models also derives from another part of the procedure used in experiments like Cummins (1995), where in a pre-test participants are asked to generate as many disablers or alternative causes as they can in one minute. For LP, presumably, this just involves reading the items off the abnormality list for the rule. But this could not be the case for the novel rules used in these experiments. For these, there are no precompiled rules with an associated abnormality lists. On our view, people build causal models of the situation described in the conditional, recruit semantic associates of those materials and check whether in their causal model the presence of the putative disabler prevents the consequent event occurring.

SL also cite evidence from autistic participants (Pijnacker, et al. 2009, 2010) for the psychological reality of abnormalities or exceptions. Of course, it is not the bare reality of

[5] As we observed, only the German suffix "ver" seemed to be a possible candidate. The old adage, "the exception proves the rule" also comes close. However, if you are like us, then when you first encountered this adage it sounded paradoxical.

exceptions against which we argued, but LP's account of how they are handled. The experiments they cite would be more compelling if LP were the only theory that predicts apparently dysfunctional exception handling and these were the only experiments demonstrating this phenomenon. However, we have shown dysfunctional exception handling in high schyzotypes (Sellen, Oaksford, & Gray, 2005). High schizotypes showed no effect of many disablers in a replication of Cummins (1995) with high and low schizotypy groups. That is, they did not show a suppression effect for MP or MT but they did show a suppression effect for AC and DA. The negative symptoms of schizophrenia are thought to be due to reduced dopamine levels in the frontal lobes. Cohen and Servan-Schreiber (1992) modelled some of the observed cognitive symptoms by turning down the gain parameter of connectionist units representing context. We modelled our findings with high schizotypes in a similar way (Oaksford & Chater, 2010b). The absence of exceptions provides the context in which a rule will work or a cause will produce its effect. Consequently, in a neural network representation of System 1, we turned down the gain of the units representing disablers, reducing their effect on inference when they were included in the premises. This one parameter change provided good fits to Sellen et al.'s (2005) results. Interestingly, this may link with a relatively old finding that dopamine levels are reduced in autism (Lake, Ziegler, & Murphy, 1977).

In this section, SL also describe our citing hash coding as an analogy for LP's abnormality mechanism as "egregiously misleading." They argue that this is because, "It is no good reacting against all immigration of formalisms from outside psychology: one simply loses the benefits of mathematics." Of course, we never argued against all immigration of formalisms from outside psychology, just LP. Indeed, we have specifically argued *for* the importation of the causal Bayes net formalism. Like SL, we fully support the need for formally specified models in reasoning research (see, Hahn, 2014). Nonetheless, we believe the hash coding analogy was apposite: both hash coding and abnormality propositions solve a computational problem in a way that seems to assume answers to the psychologically interesting questions. In the case of LP, the psychologically interesting questions are largely epistemic—e.g., how do we construct an abnormality list? Specifically, we doubt that LP solves the epistemological frame problem (Fodor, 1983, 2001; Shanahan, 2009).

In the closing paragraph of this section, SL discuss their view of the relationship between System 1 and System 2. We are in agreement with much of what they have to say here. We (Oaksford & Chater, 2013; in press) have also argued that people construct small scale models of the premises of arguments (integrated with world knowledge) from which they "read off" putative conclusions.[6] The details are beyond of the scope of this reply but we can highlight the contrasts with SL. We envisage these models as causal models where associative knowledge (represented in connectionist networks) may be used elaboratively to suggest further structure (in addition to that conveyed by the premises) and the values (or distributions of values) of the various parameters. Often, in order to make a response, more than one interrogation of these models is required (whether alternative models are considered or not). We have argued that the need to verbalise the results of the records of these interrogations may lead to systematic errors (Oaksford & Chater, in press). But these errors would be at the System 2 level. The contrast with SL is clear, where we envisage System 1 as associative and connectionist, they view it as provided by a symbolic LP data base. Moreover, where we envisage the small scale models as similar to CBNs, they view them as provided by localist[7] neural networks. The inversion of System 1 process proposed by SL in

[6] Of course, this is hardly a novel proposal (e.g., Johnson-Laird, 1983) although there is now active debate about the nature of these small scale models of the world which people construct in order to draw inferences.
[7] Localist NNs should not be conflated with local inference. Localist NNs assume non-distributed representations of object and events, e.g., there is one unit to represent say cat.

what we called in the introduction "the received view" of the System1/System 2 distinction could not be clearer.[8]

Replies to Specific Criticisms.

In the main section of the paper, SL respond to what they take to be our principle criticisms of LP. We address each of their responses in turn (SL's numbering is shown in brackets).

Inconsistency (1). In this section, SL query our specific astronomical example for being "exceptional" and then, after an observation about the representation of "theoretical and observational conditionals,"[9] they outline a variety of ways in which LP may handle inconsistencies. They note, anecdotally, that there are similarities with human experience. While this is an interesting discussion, it sidesteps the fact that we discussed the response to inconsistency in three different sections of our paper.

In one of these sections, we discussed how perceived inconsistencies can be treated as an occasion to learn, i.e., to revise the strength of a perceived dependency between antecedent and consequent. Moreover, we have shown how this strategy can provide much better fits to the experimental data on abstract conditional inference tasks (Oaksford & Chater, 2007; 2013). These data consisted of 65 individual studies using the conditional inference task involving 2774 participants. That is, this approach to inconsistency, which is not available to LP, in which there is no direct analogue of altering strength parameters,[10] provide good fits to a large amount of data. This is in contrast to LP, where there is no more than anecdotal support for its mechanisms for handling inconsistency and for which no detailed model fits have, to our knowledge, so far been provided.

[8] We also agree with SL that a lot of the action is at the System 1/model interface and that System 2, involving language and handling alternatives, remains very underspecified in dual systems approaches.

[9] This comment was rather obscure and we would value seeing it elaborated.

[10] In the noisy–OR rule, the strength of a dependency is treated as inversely proportional to the number of disablers, which is partially consistent with the approach SL discuss later. However, in the noisy–OR representation a disabler's disabling strength is also taken in to account, which SL's approach cannot deal with.

SL (p. 5) further suggest that the Bayesian probabilistic approach “does not account for inconsistency handling as much as blur it.” Yet there is good evidence that people are a sensitive to this “blurred” notion (Cruz, Baratgin, Oaksford, & Over, 2015). According to a probabilistic account based on what Edgington (1995) called the Equation, that is, Pr(*if p then q*) = Pr($q|p$), the probabilities of the premises of an inference place coherence constraints on the range of values that the probability of the conclusion can take. Moreover, these coherent probability intervals for the conclusion are a deductive consequence of the premises and their probabilistic interpretation. It has been shown that peoples’ assignments of conclusion probabilities respect probabilistic coherence (Cruz et al., 2015). This is not the same as showing that people are sensitive to violations of coherence but at least one possible reaction to detecting such a violation is clear from our learning example. They need to revise one of their premise probabilities. This is an interesting area for future investigation; indeed, we are experimentally investigating people’s reaction to coherence violations currently.

Relevance (2). We suggested that in addressing the epistemological frame problem LP would need to address the computation of relevance in System 1. We mentioned also that context occasionally helps to disambiguate relevant interpretations. SL respond by arguing that LP has been directly applied to computing contextual relevance. On this point, we would concede that LP may well be making some theoretical advances. Our own view on these issues is that mechanisms like online schema computation in constraint satisfaction neural networks (Rumelhart, Smolensky, McClelland, & Hinton, 1988) probably underpin the computation of relevance. Moreover, in the papers SL cite elsewhere (e.g., Baggio, Choma, van Lambalgen, & Hagoort, 2010, p. 2137) discussing the ERP evidence on how LP computes contextual relevance, it is conceded that, “…the ERP data reported here do *not* favour unification [the LP approach] over alternative accounts…” (our italics). So the evidence does not seem to come down unequivocally on LP’s side. Moreover, in a constraint satisfaction network, when there is no disambiguating contextual information available, the strength of connections will make different interpretations more available than others. As we indicated above, we doubt that the solution LP provides to the narrow, technical frame problem can account for the broader epistemological frame problem. This would require an account of analogy and metaphor (Shanahan, 2009) which has been argued to be the core of the unconscious reasoning system (Lakoff & Johnson, 1980, 1999). In this respect, connectionist models of System 1 may be in better shape (e.g., Thomas & Mareschal, 2001).

Enablers vs. alternative causes (3). Distinguishing the status of premises in the suppression task seems not to yield any real disagreement between us and SL. SL concede that these distinctions require general knowledge. However, the proposal that enablers will be represented as conjunctive antecedents ($p \wedge s \wedge ab' \rightarrow q$) does not seem to distinguish causes (p) from enablers (s). In contrast, there is a detailed causal model theory of these distinctions (Sloman, Barbey, & Hotaling, 2009). That is, we believe that this distinction is *not* "one that is as problematic, or not, in BNs as LP nets."

The implicit suppression paradigm (4). In this section, SL respond to our suggestion that the implicit suppression paradigm (e.g. Cummins, 1995) is problematic for LP. They argue that this paradigm shows that an alternative, more qualitative, LP approach to suppression effects is possible. This initial argument presupposes that there is a good correlation between the magnitude of suppression effects and the number of defeaters or alternative causes participants can generate in a pre-test. The pre-test in the implicit paradigm is usually carried out with a separate group of participants and SL do not cite any literature on the studies demonstrating the extent of this correlation. In, for example, Geiger and Oberauer (2007), number of disablers was an independent variable in a factorial design.[11] We do not doubt that the correlations SL mention will be observed but the upshot of Geiger and

[11] In the similar factorial designs used by Cummins (1995), effect sizes were not reported so we cannot determine whether in these studies number of different disablers "..explains[s] a good deal of variance.." or not. In Geiger and Oberauer (2007) it was the lack of any such effect that was the upshot of the paper.

Oberauer’s studies was that information about the number of different disablers is completely overridden when frequency information is available.

SL go on to propose that frequency information can be generated from the operation of LP—that is, they suggest that such information can be gleaned from intensional knowledge. We agree that this is a potentially important development. However, there are two points to make. First, the proposed mechanism apparently involves a neuron in the neural network implementation counting the times a rule/abnormality proposition is activated. This is an ad hoc addition to the proposed implementation of LP. Since it is not inherent to the LP approach it could just as well be implemented in a similar ad hoc manner in other approaches if required. Second, SL provide no reference for this work so that it can be judged independently, so it is not clear how far this approach has been developed. Nonetheless, we assume that the discussion here is intended to feed forward to SL’s discussion of Martignon et al. (in prep).

Before discussing some further aspects of Martignon et al. (in prep), SL diverge in to two long quotations from Pearl about the need for additional sources of information in order to extract causal information from raw data. This was in response to our argument concerning abnormality lists that what is normal or abnormal must be learnt from the statistical structure of the world. As we pointed out in the first section of this paper, we have argued that causation is basic. Our collective intuitions about causation from many sources allow us to form hypotheses about the structure of the world which statistical evidence can confirm or deny. This knowledge allows us to form causal hypotheses captured in the structure of a causal Bayes net. However, a great deal of what we are told we just take to be true without question, if it coheres with what we already believe. Bayesian accounts provide a formal measure of coherence (Bovens & Hartmann, 2003) for which there is some experimental support (Harris & Hahn, 2009). In sum, our argument was that what is normal and what is abnormal has to be learned; and it is learned by discovering, *by whatever means*, the causal structure of our world. As Sternberg and McClelland (2010) have recently showed, we can learn the nature and efficacy of causes either inferentially or by pathway strengthening *or*, of course, we can just accept the testimony of a reliable informant (Hahn, Harris, & Corner, 2009; Hahn, Oaksford, & Harris, 2013; Oaksford & Hahn, 2013).

In the final paragraphs of this section, SL mention the experimental work (Martignon et al., in prep) which underpins their earlier comments. They report experiments using a within subjects design. They found that ratings of the likelihood of disablers explains additional variance in predicting the inferences participants endorse, e.g., MP, over and above the sheer number of disablers generated. This experiment appears to replicate Fernbach and Erb (2013) where participants estimated disabling strength and the base rate of disablers to compute causal power in a noisy-OR gate. For each retrieved disabler, the disabling probability can then be calculated as disabling strength × base rate. Causal power (Cheng, 1997) is equal to 1 minus the aggregate of the disabling probabilities (see, Ferbach & Erb, 2013, Eqs. 3, 4, & 5; Oaksford & Chater, in press, Eqs. 8 & 9). So causal power as a predictor of endorsements of MP or MT captures both the number of disablers and what SL refer to as likelihood. Consequently, the data SL describe would seem unable distinguish between their ad hoc extension of LP and causal Bayes nets. Both seem to predict that number of disablers and disabling probability ("likelihood") will affect MP and MT ratings or endorsements.

Explaining away alternative causes (5). In this section, SL suggest that LP can explain discounting inferences in contradistinction to the argument in our original paper. However, they have misrepresented these inferences. Using their formal example, discounting occurs when comparing two situations ("*r?*" is the question posed to participants in the task):

$$p \rightarrow q, r \rightarrow q, q, r? \qquad (3)$$

$p \rightarrow q, r \rightarrow q, q, p, r?$ (4)

For example, where *p*: it is raining, *r*: the sprinklers are on, and *q*: the pavements are wet, learning that that the pavements are wet (1) makes it quite likely that the sprinklers are on (*r*). So, when asked "are the sprinklers on?" (*r?*), participants should respond that this is highly probable. However, finding out that it is raining *discounts* the sprinklers being on as the explanation of the pavements being wet—the rain alone explains why the pavements are wet, so that the mere wetness of the pavements provides no indication of whether the sprinklers were on or not. In the experimental data, both children (Ali, et al., 2010) and adults (Ali, et al., 2011) rated *r* as less likely in (4) than in (3). But SL observe that in LP for both (3) and (4), *r* is not derivable. Consequently, people should not endorse *r* given either set of premises so there should be no differences between conditions. The case SL consider is not the actual discounting inference involving the contrast between (3) and (4) but the contrast between (3') and (4):

$r \rightarrow q, q, r?$ (3')

In (3') *r* is derivable, in a way that it is not in (3). All the experimental tasks in Ali et al (2010, 2011) used the premise sets in (3) and (4). Consequently, SL fail to show how LP can account for the discounting inferences observed in these experiments. They suggest the possibility that in (3) the disjunction $p \vee r$ might be derivable at the System 2 level, that is, *outside of LP.* But to see a reduction in endorsements in (4) would seem to require an inference of the form $p \vee r, p$ therefore *not-r*, which would require the further ad hoc assumption that "∨" is interpreted exclusively.

(3) and (4) are the context in which the discounting inference has been described in the normative literature. Discounting follows as a consequence of the Markov condition in causal Bayes nets; without a similarly strong structural assumption it is not clear how LP can explain this inference pattern. Violations of this condition have been observed but these have

largely been explained within the causal Bayes net framework (Rehder, 2014). In sum, discounting is indeed a novel prediction of the Bayesian approach (Chater et al., 2011) one which we suggest that SL *cannot* explain within LP—unspecified System 2 inferential mechanisms must be invoked.

Conclusion

In concluding, we first observe that SL do not address the closing sections of our paper. We then raise some issues about where this debate could go in the future. Finally, we consider the where the burden of proof lies in this debate.

Our speculative closing remarks were aimed at providing an answer to the practical question of when in real life do we need consistent,[12] retrievable, default hierarchies? We suggested that (1) consistency is probably a local issue about the particular models we build to guide our actions in the world and not to be expected of our global world knowledge; and that (2) knowledge about most situations is shallow which is remediated by the fact that knowledge is socially distributed. We could add here that as we (Hahn & Oaksford, 2007) and others (Mercier & Sperber, 2011) have argued reasoning is engaged primarily in the service of argumentation, that is, in resolving disagreements between people. Thus, when we make public commitments we are often required to argue for their consistency with our other beliefs or the facts. Making a public commitment could also be to act on our beliefs, with inconsistency becoming apparent if those actions fail. It is only in these behavioural and social settings that consistency becomes a serious issue. Consequently, our individual global belief systems may not need to satisfy such exacting standards as global consistency. This is a potentially deflationary approach to the problem that LP tries to solve and perhaps to the frame problem in general—that is, in psychologically real belief systems this problem doesn't

[12] In LP, we assume "consistency" is not classical consistency (for any proposition p, p & $\neg p$ is not derivable) but some three valued alternative—the underlying logic is apparently Kleene's strong three valued system (Haack, 1975).

actually exist.[13] From the web of soft constraints that makes up our associative world knowledge we construct small scale models that provide the content of our thoughts and the basis for our actions the contents of which we can verbalise in language and communicate to others; but it is only at this final stage that consistency and resolving consistency is an issue and it is a local issue about a particular model.

In furthering the debate between LP and the probabilistic approach based on causal Bayes nets or probability logic there are a couple of issues worth mentioning. First, both suggest a three valued system, for LP it is Kleene's strong three valued system (Haack, 1974; Stenning & van Lambalgen, 2005) and for probability logic it is de Finetti's three valued system (Baratgin, Over, & Politzer, 2013). These systems may differ on their interpretation of the third truth value (Haack, 1974). However, the actual truth tables are identical (Baratgin et al., 2013). This means that at this level of abstraction from continuous valued probabilities (which SL don't like) there may be no grounds to distinguish LP from probabilistic approaches. It is only when we look at probabilistic inferences, like discounting, that differences emerge. . Second, at a deep theoretical level there is a close connection between Bayes nets and LP (Kowalski, 2010; Poole, 1993, p. 81)—"any probabilistic knowledge representable in a discrete Bayesian belief network can be represented in…a simple framework for Horn-clause abduction, with probabilities associated with hypotheses." Horn-clause abduction is the inferential scheme employed in LP. Working out these relationships may provide for some further productive dialogue.

In their closing paragraph, SL (p.??) place the burden of proof, "on those who would invoke probability…to show that it can do something that crude intensionally derived conditional frequencies of inference cannot." The short response to this suggestion is "explain some of the empirical data." The paper reporting this new mechanism has not been

[13] We say "potentially" because the proposal and its consequences remain to be full worked through.

published, so it cannot yet be judged whether it can be applied to the range of experimental results to which probabilistic approaches have been applied. So, the burden of proof lies with LP to demonstrate that it is capable of explaining the large range novel results spawned by probabilistic approaches.[14]

In conclusion, we thank SL for their constructive response to our critique of LP. We think many issues have been clarified. On the issue of dual systems there are points of agreement—the System 1/model interface and the current underspecification of System 2—but also important areas of disagreement—we cleave much closer to the received view of the System 1/System 2 distinction. The fundamental disagreements surround the treatment of nonmonotonicity and uncertainty and whether this only a local or global problem for real human reasoners. As with the comparison between all theories, eventually the data will be the ultimate determiner of which approach is correct and in this respect we suggest that probabilistic approaches seem to have the edge, at least for the moment.

[14] We spare the reader a very long list of citations establishing this claim but rather refer the reader to the most recent textbook on thinking and reasoning (Manktelow, 2012) where the new paradigm features heavily.

References

Adams, E. W. (1998). *A primer of probability logic*. Stanford: CLSI Publications.

Ali, N., Chater, N., & Oaksford, M. (2011). The mental representation of causal conditional inference: Causal models or mental models. *Cognition*, *119*, 403-418. doi: 10.1016/j.cognition.2011.02.005.

Ali, N, Schlottmann, A., Shaw, C., Chater, N., & Oaksford, M., (2010). Conditionals and causal discounting in children. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals*: *Probability and logic in human thinking* (pp. 117-134). Oxford: Oxford University Press.

Baggio, G., Choma, T., van Lambalgen, M., and Hagoort, P. (2010). Coercion and compositionality: an ERP study. *Journal of Cognitive Neuroscience*, *22*, 2131–2140.

Baratgin, J., Over, D. E., & Politzer, G. (2013). Uncertainty and the de Finetti tables. *Thinking & Reasoning*, *19*, 308-328. doi:10.1080/13546783.2013.809018

Bovens, L, & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.

Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 61-83. doi:10.1016/0010-0277(89)90018-8

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Chater, N., Goodman, N., Griffiths, T., Kemp, C., Oaksford, M., & Tenenbaum, J. (2011). The imaginary fundamentalists: The unshocking truth about Bayesian cognitive science. *Behavioral & Brain Sciences*, *34*, 194-196.

Chater, N., & Oaksford, M. (2006). Mental mechanisms. In K. Fiedler, & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 210-236). Cambridge: Cambridge University Press.

Chater, N., & Oaksford, M. (Eds.) (2008). *The probabilistic mind*: *Prospects for Bayesian* *cognitive science*. Oxford: Oxford University Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405. doi:10.1037/0033-295X.104.2.367

Chisolm, R. (1946). The contrary-to-fact conditional. *Mind*, *55*, 289-307.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, *99*, 45-77. doi:10.1037/0033-295X.99.1.45

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Reasoning with if and ands and ors. *Frontiers in Psychology*, *6*, 192. doi: 10.3389/fpsyg.2015.00192

Cummins, D. D. (1995). Naïve theories and causal deduction. *Memory & Cognition*, *23*, 646-658.

Douven, I., & Verbrugge, S. (2013). The probabilities of conditionals revisited. *Cognitive Science*, doi:10.1111/cogs.12025

Edgington, D. (1995). On conditionals. *Mind, 104*, 235–329.

Evans, J. St.B. T. (2009). *In two minds*: *Dual processes and beyond*. Oxford: Oxford University Press.

Fernbach, P. M., & Erb, C. D. (2013). A quantitative causal model theory of conditional reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1327-1343. doi:10.1037/a0031851

Fodor, J. A (1983). *The modularity of mind*. Cambridge: MIT Press.

Fodor, J. A (2001). *The mind doesn't work that way*. Cambridge: MIT Press.

Fugard, J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 635-648.

Geiger, S. M., & Oberauer, K. (2007). Reasoning with conditionals: Does every counterexample count? It's frequency that counts. *Memory & Cognition, 35*, 2060-2074.

Gilio, A., & Over, D. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, *56*, 118-131. doi:10.1016/j.jmp.2012.02.006

Goodman, N. (1947). The problem of counterfactual conditionals. *Journal of Philosophy*, *44*, 113-128.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Haack, S. (1974). *Deviant logic*. Cambridge: Cambridge University Press.

Hahn, U. (2014). The Bayesian boom: good thing or bad? *Frontiers in Psychology*, *5*, 765. doi: 10.3389/fpsyg.2014.00765

Hahn, U., Harris, A.J.L., & Corner, A.J. (2009). Argument content and argument source: An exploration. *Informal Logic*, *29*, 337-367.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*, 704-732.

Hahn, U., Oaksford, M. & Harris, A.J.L. (2013). Testimony and argument: A Bayesian perspective. In F. Zenker (Ed.). *Bayesian argumentation* (pp. 15- 38). Dordrecht: Springer.

Harris, A. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1366-1373. doi:10.1037/a0016567

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Kowalski, R. (2010). Reasoning with conditionals in artificial intelligence. In M. Oaksford, N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 254-282). New York, NY US: Oxford University Press.

Lake, C., Ziegler, M.G., & Murphy, D.L. (1977). Increased norepinephrine levels and decreased dopamine-β-hydroxylase activity in primary autism. *Archives of General Psychiatry*, *34*, 553-556. doi:10.1001/archpsyc.1977.01770170063005.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. New York: Basic Books.

Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Manktelow, K. I. (2012). *Thinking and reasoning*. Hove, East Sussex: Psychology Press.

Martignon, L., Stenning, K., & van Lambalgen, M. (in prep.). Qualitative models of reasoning and judgment: Theoretical analysis and experimental exploration.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57-74. doi:10.1017/S0140525X10000968

Oaksford, M. & Chater, N. (Eds.) (1998). *Rational models of cognition*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2010a). Causation and conditionals in the cognitive science of human reasoning. *Open Psychology Journal*, *3*, 105-118. (Special issue on *Causal learning beyond causal judgment*, Edited by J. C. Perales, & D. R. Shanks).

Oaksford, M., & Chater, N. (2010b). Conditionals and constraint satisfaction: Reconciling mental models and the probabilistic approach? In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals*: *Probability and logic in human thinking* (pp. 309-334). Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking and Reasoning, 19*, 346-379. doi:10.1080/13546783.2013.808163

Oaksford, M., & Chater, N. (2014). Probabilistic single function dual process theory and logic programming as approaches to non-monotonicity in human vs. artificial reasoning. *Thinking and Reasoning*, *20*, 269-295. doi: 10.1080/13546783.2013.877401

Oaksford, M., & Chater, N. (in press). Causal models and conditional reasoning. In M. Waldmann (Ed.), *Oxford handbook of causal cognition*, Oxford: Oxford University Press.

Oaksford, M., & Hahn, U. (2013). Why are we convinced by the ad hominem argument?: Bayesian source reliability and pragma-dialectical discussion rules. In F. Zenker (Ed.). *Bayesian argumentation* (pp. 39-59). Dordrecht: Springer.

Over. D. (2009). New paradigm psychology of reasoning. *Thinking and Reasoning*, *15*, 431-438.

Over, D. E., Hadjichristidis, C., Evans, J. St.B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*, 62-97. doi:10.1016/j.cogpsych.2006.05.002

Nute, D. (1984). Conditional logic. In D. Gabbay, and F. Guenthner (Eds.), *Handbook of philosophical logic*, Vol. 11 (pp. 387-439). Dordrecht: D. Reidel Publishing Company.

Pfeifer, N., & Kleiter, G. (2010). Mental probability logic. In M. Oaksford, M. and N. Chater (Eds.). *Cognition and conditionals: Probability and logic in human thinking* (pp. 153-

173). Oxford, UK: Oxford University Press.

Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J., & Hagoort, P. (2010). Reasoning with exceptions: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, *23*, 471–480.

Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J., Kan, C., & Hagoort, P. (2009). Conditional reasoning in high-functioning adults with autism: Evidence for impaired exception-handling. *Neuropsychologia*, 47, 644–651.

Poole, D. (1993). Probabilistic Horn abduction and Bayesian networks. *Artificial intelligence*, *64*, 81-129.

Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, *72*, 54-107. doi:10.1016/j.cogpsych.2014.02.002

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1988). Schemata and sequential thought processes in PDP models. In A. M. Collins, E. E. Smith, A. M. Collins, E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 224-249). San Mateo, CA, US: Morgan Kaufmann.

Sellen, J., Oaksford, M., & Gray, N. (2005). Schizotypy and conditional inference. *Schizophrenia Bulletin*, *31*, 105-116.

Shanahan, M. (2009). The frame problem. Retrieved December 12, 2015, from http://plato.stanford.edu/entries/frame-problem/

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

Sloman, S., Barbey, A. K., & Hotaling, J. M. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science*, *33*, 21-50. doi:10.1111/j.1551-6709.2008.01002.x

Sowa, J. F. (2000). Processes and causality. Retrieved September 30, 2015, from http://www.jfsowa.com/ontology/causal.htm

Stalnaker, R.C. (1968). A theory of conditionals. In N. Rescher (Eds.), *Studies in logical theory* (pp. 98-112). *American Philosophical Quarterly Monograph Series*, *2*. Oxford: Blackwell.

Stalnaker, R. C. (1984). *Inquiry*. Cambridge, MA: MIT Press.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. Oxford: Oxford University Press.

Stenning, K., & van Lambalgen, M. (2005). Semantic interpretation as computation in nonmonotonic logic: The real meaning of the suppression task. *Cognitive Science*, *29*, 919-960. doi:10.1207/s15516709cog0000_36

Sternberg, D. A., & McClelland, J. L. (2012). Two mechanisms of human contingency learning. *Psychological Science*, *23*(1), 59-68

Thomas, M.S.C., & Mareschal, D. (2001) Metaphor as categorisation: a connectionist implementation. In J.A. Barnden, & M.G. Lee, (Eds.) *Metaphor and artificial intelligence* (pp. 5-27). Metaphor & Symbol 16. London, UK: Psychology Press.

van Lambalgen, M., & Hamm, F. (2004). *The proper treatment of events*. Blackwell, Oxford and Boston.