Causal Discounting and Conditional Reasoning in Children

Nilufa Ali
Cognitive, Perception and Brain Sciences
University College London
Anne Schlottmann
Developmental Science
University College London
Abigail Shaw
Developmental Science
University College London
Nick Chater
Cognitive, Perception and Brain Sciences
University College London
Mike Oaksford
School of Psychology
Birkbeck College London

1. Introduction

Conditional reasoning is widely viewed as central to human inference and a paradigm case of logical thought. Perhaps the most natural use of the indicative conditional is to express relationships between cause and effect, either in reasoning predictively from causes to effects (if I cut my finger, then I bleed); or reasoning diagnostically from effects to causes (if I bleed, then I cut my finger). To the extent that conditionals are used to describe cause-effect relationships, the study of cause and effect may help illuminate the function of conditionals.

Many studies of conditional reasoning have used experimental materials which participants are likely to interpret in causal terms. Indeed, unless the domain of reasoning is highly abstract (e.g., about numbers and letters printed on cards), it seems inevitable that people will recruit causal knowledge (about fingers, cuts, bleeding, etc.) to solving conditional reasoning problems. But in the majority of studies causal structure has been incidental, rather than the focus of inquiry (but see Griffiths & Tenenbaum, 2005; Sloman & Lagnado, 2005). In this chapter we suggest that the causal structure described by the premises may be more important than its logical structure—the traditional focus of research in the psychology of conditional reasoning—in determining the inferences people draw.

More specifically, we consider that conditional inference patterns should differ when reasoning from cause-to-effect (predictively) than when reasoning from effect-to-cause (diagnostically). Accordingly, the present approach will develop a probabilistic view of indicative reasoning about cause and effect. There have been many studies marshalling evidence for a probabilistic approach to understanding conditionals, (e.g.,

But if participants interpret the premises as causal, then a theory of causal inference may fare better.

In fact, an influential normative account of causal reasoning recently developed in artificial intelligence, (Pearl, 2000) allows us to make predictions about the effects of causal direction not made by existing conditional reasoning theories. Our previous research with adults (Ali, Chater, & Oaksford, 2008) broadly confirms the predictions of this account. However, there was an anomaly that we interpreted as being caused by the experience of our adult participants and which may be absent in children with less experience with causal relations. Accordingly, the aim of the experiment reported here was to investigate the influence of causal direction on conditional reasoning in children.

We first outline the conditional reasoning paradigm and the standard results. We then describe Pearl’s (2000) normative theory from which we develop the current account.

**Conditional Inference**

The starting point for the psychology of conditional inference has been the formal analysis of the conditional provided by standard logic. According to standard logic, a conditional if \( p \) then \( q \) is true, and so can be accepted, if and only if either the antecedent \( p \) is false or the consequent \( q \) is true. This semantics for the conditional licenses two formal rules of inference called modus ponens (MP) and modus tollens (MT):

\[
\text{MP} \quad \frac{p \rightarrow q, p}{\therefore q} \quad \text{MT} \quad \frac{p \rightarrow q, \neg q}{\therefore \neg p}
\]

These inference schemata read that if the propositions above the line are true, then it can be inferred that the propositions below the line are true. For example, for MP, if it is true that if John has a runny nose \( (p) \), he has a cold \( (q) \) and that John has a runny nose, then it is true that Johnny has a cold. According to standard logic both MP and MT inferences are valid. Consequently, if people are logical then they should endorse both inferences and reject the inferential fallacies of denying the antecedent (DA) and affirming the consequent (AC):

\[
\text{DA} \quad \frac{p \rightarrow q, \neg q}{\therefore \neg p} \quad \text{AC} \quad \frac{p \rightarrow q}{\therefore q}
\]

However, rather than exhibiting the predicted symmetry between MP and MT, participants tend to endorse MP far more than MT. For example, in a recent meta-analysis involving 65 conditional inference experiments and 2774 participants, 97% (SD = 3.6%) on average drew the MP inference but only 72% (SD = 15.5%) the MT inference (Schroyens & Schaeken, 2003, see also Oaksford & Chater, 2003). Over the 65 studies, this result represents a highly significant asymmetry, \( t(64) = 15.44 \), \( p < .0001 \), between MP and MT. Moreover, participants also endorse DA and AC and there is a similar asymmetry such that AC is drawn more than DA.

**Causal Conditional Reasoning**
The results we described in the last section almost all relate to abstract alphanumeric stimuli rather than the causal conditionals that are the focus of the current study. However, there has been much work investigating the effects of causal conditionals on people’s reasoning. Causal conditionals, in particular, have been used to show that the inferences, MP and MT, and the fallacies, DA and AC, can be suppressed by providing information about possible defaults, i.e. about additional enabling conditions or alternative causes. For example, if you are told that if the key is turned the car starts and that the key is turned, you are likely to endorse the MP inference to the conclusion that the car starts. However, if you are also told that the petrol tank is empty, you are less likely to endorse this conclusion, because the car will not start if the petrol tank is empty.

A full petrol tank provides an enabling condition that allows turning the key to start the car. The petrol tank being empty will also affect MT. If you knew that the car didn’t start you may not infer that the key was not turned because the empty petrol tank may be the cause of the car not starting.

Information about alternative causes can suppress DA and AC. For example, if you are told that if the key is turned the car starts and that the key is not turned, you might endorse the DA inference to the conclusion that the car does not start. However, if you are also told that the car was hot-wired, you may be less likely to endorse this conclusion because the car may start even though the key was not turned because it has been hot-wired. This alternative cause would also mean that you are less likely to endorse AC. If you knew that the car started you may not infer that the key was turned because the car starting may have been caused by being hot-wired.

These kinds of effects have been investigated using two paradigms, the explicit and the implicit paradigms. We describe each, as features of both will figure in the design of our experiment. Although Byrne (1989) did not originally use causal materials she demonstrated all the above effects by providing participants with additional conditional statements containing the new information:

<table>
<thead>
<tr>
<th>Enabling Conditions (affecting MP)</th>
<th>Alternative Causes (affecting AC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>If the key is turned the car starts</td>
<td>If the key is turned the car starts</td>
</tr>
<tr>
<td>If there is fuel in the tank the car starts</td>
<td>If it is hot-wired the car starts</td>
</tr>
<tr>
<td>The key is turned</td>
<td>The car starts</td>
</tr>
<tr>
<td>The car starts?</td>
<td>The key was turned?</td>
</tr>
</tbody>
</table>

Cummins, Luhars, Alkonis, and Rist (1991, see also Cummins, 1995) and Thompson (1994) report results that were very similar to Byrne (1989). However, they left information about additional and alternative antecedents implicit. That is, unlike Byrne (1989), these authors pre-tested their conditional rules for how many alternative and additional antecedents came to mind then they used these rules in the experimental task without explicit cuing concerning the relevance of alternative and additional antecedents. Cummins (1995) also used diagnostic conditionals, e.g., if the car starts then the key was turned, and showed a reversal of these effects. That is, new alternative causes affected MP and MT and enabling conditions affected AC and DA. Such effects are not predicted by logic and demonstrate clear effects of causal factors on conditional reasoning; although it has been argued that such effects can be incorporated in logic based models (see Byrne, Espino and Santamaria, 1999).
Pearl's Normative Theory and Predictions

Pearl’s (2000) normative theory of causal reasoning provides a rich conceptual and mathematical framework for dealing with a wide range of issues in causal reasoning. Here we will focus on its broad structure, and its key predictions.

Pearl’s account shows how causal relationships can be represented in terms of graphical structures (see also Pearl, 1988), such as those in the second row of Table 1. The structure for condition 1 and 3 shows two possible causes, C1 and C2, of a common effect, E. The structure for condition 2 and 4 shows a common cause, C, of two effects, E1 and E2. The direction of an arrow represents the direction of causality. The “nodes,” which are connected by the arrows, represent variables, which interact causally. In these experiments, these two structures are described using causal and diagnostic conditionals. So the common effect structure is described as if C1 then E and if C2 then E (causal conditionals in CE direction) and the common cause structure is described as if E1 then C and if E2 then C (diagnostic conditionals in EC direction) (see Cummins, 1995). In other words, the cause is either stated as antecedent or consequent of the conditional.

Table 1 about here

Given these descriptions, which form two premises in the inferences we investigate, there are four inferential patterns shown in Table 1. They arise from two binary factors: direction (cause-to-effect or effect-to-cause); and whether the consequent of the conditional is added as an additional premise or not (this is explained shortly). Rather than trace through the formal predictions of Pearl’s theory here, we focus on the task given to the participant, and the rationale for the pattern of results that would be expected if participants are following Pearl’s (2000) normative account.

We first consider the Non-Consequent condition in the left columns of Table 1. Participants are first given the two causal conditional premises (“background” information, B), then asked to rate the probability of the first antecedent (Rating 1). Participants are then told that the second antecedent is true. What we are interested in is whether the additional information about the second antecedent raises or lowers the rated probability of the first antecedent or leaves it unchanged. Note that in conditions 1 and 2, for the cause-to-effect and effect-to-cause case, the problem has exactly the same logical form, but according to Pearl’s causal analysis, these cases are very different.

In the cause-to-effect case, the key question is whether the addition of the new knowledge of one cause (C2) has an impact on the rated probability of another cause (C1), given the background knowledge that both causes give rise to effect E. We do not know whether E has occurred, so the first rating of C1, whether it rains, ought to simply reflect the base rate of the event. When we then learn that C2 occurred, i.e., that the sprinklers are on, this should have no effect on the second probability rating for C1, because rain and sprinklers are independent of each other (at least if the sprinklers are controlled automatically). Thus the occurrence of one event has no impact on the occurrence of the other. Formally, in terms of Pearl’s theory, the only path between the two causes (see Table 1) passes through a “collider” where the two arrows meet at a single node, thus Prob(C1|B) = Prob(C1|C2, B).

In contrast, consider the effect-to-cause case of condition 2. Here, knowing one effect (that there are shadows) of it being sunny should raise the probability of other effects (that it is warm outside). This is because it being sunny causes both shadows and warmth, and hence the presence of one effect is informative about the presence of
another. This is a case of causal augmentation. In terms of Pearl’s theory, the two effects are joined by a path that does not include a “collider” and thus \( \text{Prob}(E_1|B) < \text{Prob}(E_1|E_2, B) \).

The pattern changes subtly, but significantly, when a further premise (the “consequent,” see the right hand columns of Table 1) is added to the two background conditional statements. In the cause-to-effect case (condition 3), Rating 2 should decrease from Rating 1. Now that we are told initially that the street is wet, we can reason that this is likely to have been caused either by rain or sprinklers. Hence, for the initial judgement, both causes are equiprobable and reasonably likely, leading to a higher Rating 1 than in condition 1. However, once one cause is confirmed (sprinklers), this suffices to explain that the street is wet, thus the effect ceases to provide evidence for rain. Hence Rating 2 should decrease, to the base rate of rain. This is a case of causal discounting. In terms of Pearl’s theory, the collider linking two causes ‘unblocks’ the path between \( C_1 \) and \( C_2 \) and thus \( \text{Prob}(C_1|E, B) > \text{Prob}(C_1|E, C, B) \).

Turning now to our final effect-to-cause case (condition 4), here the addition of the consequent (the cause, “it is sunny”) has a different result. Without the consequent, in condition 2, learning about one effect (shadows) increases the probability of another (warmth), because one effect increases the probability that the cause was present which in turn makes other effects more probable. But if the cause is given initially, then Rating 1 should already be at ceiling, and cannot be raised further by learning about the other effect. That is, the path between \( E_1 \) and \( E_2 \) is “blocked” by the value of \( C \) being specified; once \( C \) is known, \( E_2 \) carries no further information about \( E_1 \), i.e., \( \text{Prob}(E_1|E_2, C, B) = \text{Prob}(E_1|E_2, C, B) \).

These predicted equalities and inequalities form the basis of our experimental predictions. Using the Rating 2 (RH side of each equality/inequality) minus Rating 1 (LH side of each equality/inequality) difference (\( \Delta R \)) as the dependent variable, the experiment is a 2 (causal direction: CE, EC) by 2 (consequent: consequent (C), non-consequent (NC)) factorial design. At the highest level we predict two main effects, effect-cause > cause-effect (EC > CE) and non-consequent > consequent (NC > C). More detailed predictions correspond to the four simple effects (EC > CE, ECNC > CENC, ECNC > ECC, CENC > CEC). The most detailed level of predictions correspond directly to the equalities and inequalities outlined above (ECNC > 0, CEC < 0, CENC = 0, ECC > 0). These predictions are shown in the interaction plot in Figure 1. These predictions show no interaction effect. However, the underlying probabilities depend on the particular contents used in the experimental materials. Consequently, the lack of an interaction is not a direct prediction as one may arise as a result of the particular materials used.

Figure 1 about here

These predictions can be derived directly from Pearl’s normative theory of causality, but could not be made simply by considering the logical form of these conditional arguments (see Sloman and Laguné [2005] for a similar argument). Ali, Clutter, and Oakford (2008: Experiment 1) tested these predictions with 10 different scenarios. The overall results are shown in Figure 2. These results show that most of the predicted effects were observed. However, there were some discrepancies. These related to conditions 1 and 4, where \( \Delta R \) is predicted to be zero. In condition 4, \( \Delta R \) was significantly lower than zero, i.e., \( \text{EC} < 0 \).
ΔR was close to being significantly lower than zero, i.e., CENC < 0 \(P(C_1|C_2, B) < P(C_1|B)\). Thus causal discounting was observed more widely than predicted by Pearl’s theory, in three rather than just one of four conditions.

**Figure 2 About Here**

We interpret this finding as relating to the effects of prior knowledge indicating independence of the effects of a common cause, or (negative) dependence of the causes of a common effect.

Take the case of the common cause structure in a diagnostic conditional. In condition 4, producing the ECC anomaly, one is told that the cause is present, i.e., it is sunny, from which one would expect both effects. However, you are then informed about only one of those effects, there are shadows. A possible pragmatic implicature is that the other effect did not occur, otherwise why explicitly mention only one of them? Another example would be "The gong was struck, it vibrated...but made no sound." That is, explicit mention of one effect may trigger the expectation that another did not occur. Moreover, people will be aware of cases in which normally positively correlated effects become uncorrelated: indoors, warmth and shadows do not have a common cause, at the poles when it is sunny there will be shadows but it will not be warm. This implicature depends on having the additional information that it is actually sunny; as long as this is not known (condition 2), mention of one effect augments the probability of the other. But when it is stated explicitly that it is sunny, implying both effects, then mention of only one of them pragmatically suggests an exception. So world knowledge -- that even if effects are normally correlated there can be exceptions -- together with conversational pragmatics leads to a reduced Rating 2.

In the common effect, causal conditional case of condition 1, producing the CENC anomaly, the events are independent and occurrence of one should not affect the other. However, although the mechanisms producing rain and those controlling the sprinklers are independent, the events can be correlated due to human intervention. In particular, some causes come into play typically when the normal cause has failed (Oaksford & Chater, 2003, 2008). Thus sprinkler systems are installed in areas where it does not rain a lot and predominantly used when it does not rain. In other words, world knowledge also suggests that physically independent causes of a common effect may in practice be negatively correlated. This would produce discounting rather than independence in the CENC condition, with a reduced Rating 2.

**Children’s Reasoning with Causal Conditionals: The Present Study.**

Recent developmental work on simple causal conditionals in children shows that suppression of the conditional fallacies occurs in children as in adults, that it improves developmentally, and that it depends on how easy it is to generate alternative cases (Barrouillet, Markovits & Quinn, 2001; Janeva-Brennan & Markovits, 1999). We also know that even pre-school children are capable of learning the more complex causal structures of the type assumed in the causal Bayes net framework (e.g., Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Schulz, Gopnik, & Glymour, 2007). These results suggest that children might also be capable of reasoning about conditionals involving common cause or effect structures, and that they might be sensitive to the factors that underpin the predictions tested by Ali et al (2008).
One might further conjecture that with less experience of the real world exceptions to
the assumptions of such structures, children might more closely follow the prescriptions
of Pearl’s theory. However, the previous explanations suggest an asymmetry between
condition 1 (CENC = 0) and condition 4 (ECC = 0). Whereas one might expect
considerable experience of conversational conventions and knowledge of causes not
stated in the problem to be required for the sophisticated pragmatic implicatures involved
in producing condition 4’s deviation from 0, condition 1 only requires a general idea of
alternative causes – one precluding the other – without additional world knowledge being
required to come up with an example. Consequently, we might expect that young
children might not show the ECC anomaly, because they do not have the necessary
knowledge, i.e., $\Delta R$ will not deviate from zero for condition 4, but we may still observe
a CENC anomaly, i.e., $\Delta R$ may still be less than zero in condition 1. The present study,
therefore, investigated how children interpret similar situations to those described in
Table 1 and whether they show the same pattern of results as predicted for Ali et al’s

Children could not be tested with purely verbal material; therefore we adapted the
task to a simpler, concrete form, where the causal structure was physically seen. The
cause-to-effect condition was represented by a box with two buttons, turning on one light
and the effect-to-cause condition was represented by a box with one button, turning on
two lights (see Table 2). The boxes were decorated with pictures illustrating two
scenarios, with two versions each for the two causal structures. In the common effect
version of the flower scenario, one cause button could represent the sun shining and the
other button could represent the presence of water. The effect light would represent a
flower, which “sparkles” if one or both buttons are pressed. In the common cause
version, the single cause button for the shining sun made both a blue flower and a red
flower light up. The other scenario concerned a ship and how the captain or engineer
could work the anchor and its light, or how the captain could work lights on the anchor
and the top of the stack. The adult study used different verbal scenarios for different
causal structures, making it difficult to separate the effects of causal structure from
scenario effects. In the child study, in contrast, the two scenarios were counterbalanced
across children, so that we may observe “pure” effects of causal structure.

Table 2 about here

In each condition, children were shown one of the boxes and then questioned in a
similar way to Ali et al (2008), but all verbal statements were accompanied by
corresponding actions with the box. For example, to illustrate the background conditional

premises,

If the sun is shining, then the flower sparkles

If there is water, then the flower sparkles,

the sun button was pressed, leading to the flower light sparkling, then the water button
was pressed, also leading to the flower light sparkling.

To vary children’s knowledge or ignorance about the state of the world/box, a
scarf covered up parts of the apparatus and only the information known at that point in
time was revealed. For example, when the initial R1 rating of the antecedent was made,
there were no additional premises and the whole apparatus was covered, so that children
did not know about the state of any of the lights/blocks.

For the second R2 rating, additional information was given, e.g.,
Now you can see that the sun is shining. How sure are you that there is water? Here, the sun button was revealed and the children saw it pressed, but the water button and flower light remained covered. In this way, the questions could be asked in the same way as in the adult study, but the verbal information was amplified visually by showing the children the corresponding state of the box.

Method

Participants. The participants were 48 children drawn from a school in Harrow, North London. Half of the children were in Year 2 (mean age = 7 yrs 1 months, range = 6 yrs 5 months to 7 yrs 9 months) and half were in Year 3 (mean age = 8 yrs 0 months, range = 7 yrs 6 months to 8 yrs 6 months).

Design. Each child saw all four conditions of interest. The CE version of one scenario was used together with the EC version of the other scenario. In each, the child first rated the non-consequent, then the consequent condition. Which scenario and which causal direction came first was counterbalanced across children. Thus the experiment was a mixed $2 \times 2 \times 2 \times 2$ design with causal direction (CE, EC) and (consequent (C), non-consequent (NC)) as within subject factors and year (Year 2, Year 3), consequent order (CE first, EC first) and scenario order (flower first, ship first) as the between subjects factors, and with the Rating 2 minus Rating 1 difference (dR) as the dependent variable.

Materials. The boxes of Figure 3 were used to illustrate different causal scenarios. They could be covered with a thick black scarf, as the problem demanded. Children responded by pointing/marking 10 cm rating bars, with shading continuously graded from dark to light. The dark end was marked "yes", the light end "no", and "not sure" appeared in the middle. Ratings were read to the nearest .5 cm.

Procedure. Participants were tested individually. Before the boxes were presented children first practiced use of the rating bar. The children were told that they would be asked how sure they were that certain things were happening and would have to answer by pointing along the scale. They were told that the surer they were that the answer was 'yes' or 'no', the further they should point to the corresponding end of the scale. If they were not sure at all they could mark in the middle.

To practice, children were asked to point to how sure they were that a white button would be picked from an envelope with different numbers of white and blue buttons inside. First the children were shown five white buttons and all children pointed to the 'yes' end. Then the children were shown five blue buttons, then three of each colour and finally five white buttons and one blue button. The next stage did not take place until the children were able to show that they understood how to use the rating scale.

The children were then shown the first box. The experimenter began explaining the causal scenario and said to the child, for instance in the cause-to-effect flower condition, “My magic box has a big flower in the middle which shimmers and sparkles. Would you like to see it sparkle?” When the children responded, the button was pressed to demonstrate its action. The experimenter then told the child that the flower only sparkles when it is happy. Pointing to the sun button and watering can button, the experimenter explained that both the sun and water will make the flower happy. Then the background conditional premises were introduced and the child was told that, “If the sun
is shining, then the flower sparkles.” and “If there is water, then the flower sparkles.” The corresponding buttons were pressed to demonstrate.

The non-consequent condition was always presented first, so children were then asked for an initial rating of the antecedent; in this example, how sure they were that there was water, without any knowledge of the state of the box. To illustrate, the box was completely covered by the scarf. Children were then told about the other antecedent, here, the alternative cause, the sun, with only the corresponding button revealed and pressed, and again asked to rate the antecedent.

The children were then tested in the same manner in the consequent condition of the same scenario, introduced as events happening on another day. In this condition, prior to their R1 rating children were told about the consequent. To illustrate, the flower was uncovered and the children saw it sparkling, whilst the cause buttons remained covered. The rest of the procedure was as before, except that the flower remained uncovered. The children were then presented with the second scenario illustrating the other causal direction and the procedure was repeated.

Results and Discussion

We first calculated ΔR for each causal direction. We then carried out a $2 \times 2 \times 2 \times 2$ mixed ANOVA with causal direction and consequent as within subjects factors and year, direction order, and scenario order as the between subjects factors, with ΔR as the dependent variable. None of the between subjects factors produced significant main effects or interacted significantly with any combination of other factors. Therefore, the counterbalancing was effective and the age differences in the sample did not affect the results. Consequently, we only report the results of a within subjects $2 \times 2$ ANOVA with causal direction (CE, EC) and consequent (consequent (C), non-consequent (NC)) as factors, and with ΔR as the dependent variable.

Figure 4 About Here

The results are shown in Figure 4. Both main effects were in the predicted directions. ΔR was significantly higher in the effect-cause direction than for the cause-effect direction (EC > CE), $F(1, 47) = 20.16, MSe = 12.28, \eta^2 = .30, p < .0001$. ΔR was also higher in the non-consequent condition than in the consequent condition (NC > C), although this effect was only marginal, $F(1, 47) = 2.41, MSe = 15.22, \eta^2 = .05, p = .063$. As Figure 4 shows, unlike Ali et al (2008: Experiment 1) there was no significant interaction effect, $F(1, 47) < 1$.

At the next level, we looked at all the simple effects using planned r-tests. In the effect-cause direction, ΔR was significantly higher in the non-consequent condition than in the consequent condition (ECNC > ECC), $t(47) = 1.73, p < .05$. However, in the cause-effect direction, ΔR was not significantly higher in the non-consequent condition than in the consequent condition (CENC > CE), $t(47) = .45, n.s$. In the non-consequent condition, ΔR was significantly higher in the effect-cause direction than in the cause-effect direction (ECNC > CENC), $t(47) = 3.59, p < .001$. Moreover, in the consequent condition, ΔR was significantly higher in the effect-cause direction than in the cause-effect direction (ECC > CEC), $t(47) = 2.26, p < .025$.

At the next level, we checked for the predicted differences from ΔR = 0, which amounts to testing the original equalities and inequalities in Table 1. In the non-consequent condition 2 (effect-cause direction), ΔR was significantly greater than zero
show that children can, on occasion, be more correct than adults, who are mislead by experience to over-apply the idea of causal discounting.

It is an open question, however, whether by pushing further down the age range, we may observe all the patterns of inference predicted by a causal Bayes approach, i.e., whether even younger children with even less experience will apply the idea of causal independence even more widely and even more appropriately. As argued above, if non-normative causal discounting depends only on specific knowledge -- of the rules of conversation, of unstated other potential causes at work in the situations described, of specific instances of negatively correlated causes -- then children may ultimately not show the anomalies found in adults.

Alternatively, it may be that causal discounting, over-generalized in adults (Ali et al., 2008) and children here, may reflect a stronger or perhaps earlier idea than independence, over-generalized from its first appearance. In particular, children might have an early idea that causes are negatively correlated, such that one cause precludes another. In fact, they might apply this interpretation not just to the CENC, but also the CEC case, where discounting would be expected for negatively correlated and for independent causes. In other words, children, and perhaps even adults, might show normative discounting for a non-normative reason in the CEC case.

In line with this view, a large literature on children’s causal attributions from the 70ies and 80ies shows that children from at least 3 years have no difficulty at all with the idea that one, but not another cause produced an effect, while the possibility of two causes never seems to be considered. In addition, Sobel, Tenenbaum and Gopnik (2004) showed backwards blocking in children as young as 3 years.

(R1 = .59, R2 = .73, ECNC > 0), F(47) = 1.89, p < .05, as for adults, and in line with the augmenting prediction based on Pearl’s theory. For condition 1 (cause-effect direction), ΔR was also significantly less than zero (R1 = .58, R2 = .44, CENC = 0), F(47) = 2.22, p < .025, i.e., children showed an anomaly, with discounting instead of independence, similar to adults in Ali et al. (2008). In the consequent condition 3 (cause-effect direction), ΔR was significantly lower than zero (R1 = .67, R2 = .49, CEC = 0), F(47) = 2.66, p < .01, as for adults, and in line with Pearl’s discounting prediction. In condition 4 (effect-cause direction), ΔR did not differ significantly from zero (R1 = .80, R2 = .80, ECC = 0), F(47) = .08, p = .47 (one-tailed). In other words, children did not show the anomaly that Ali et al. (2008) found for adults, and behaved in line with the independence prediction of Pearl’s theory instead.

In summary, at all levels of statistical analysis the majority of the predictions of the hypothesis that children take causal knowledge into account in conditional inference have been confirmed. As expected, the ECC anomaly for condition 4 observed by Ali et al (2008) was removed by moving down the age range. This result suggests that, at age 7 to 8, children are yet to acquire the knowledge required for the sophisticated pragmatic implicatures that may underpin this case of non-normative causal discounting in adults. However, children at this age continue to reveal the CENC anomaly.

The ECC results found here suggest that children of primary school age appreciate the idea of causal independence to the extent that they will apply it correctly in some circumstances in which adults’ pragmatic knowledge misleads them. This is a novel finding in the literature on children’s causal understanding. Gopnik and colleagues (e.g., Gopnik et al., 2003; Schulz et al., 2007) have shown that even preschoolers can infer causality from patterns of conditional dependence and independence, in line with Pearl’s theory, in simple cases that seem equally intuitive to adults and children. Our finding goes beyond this to
A similar debate also exists in the literature on children’s understanding of logical connectives. The traditional view here is that children initially interpret the connective “or” exclusively, with the inclusive interpretation only achieved several years later (Braune & Rumain, 1981; 1983). However, this view has recently been challenged (see Crain et al., 2000). These authors argue that even young children are capable of understanding that ‘A or B’ can refer to either or both items, and that children assume the exclusive interpretation only as a result of pragmatic implication. Further empirical work is clearly necessary to decide whether equivalent pragmatic factors are at work in the case of causal discounting.

In evaluating the wider import of this work we now assess whether and how Mental Models and Causal Bayes theories of reasoning can explain the children’s pattern of results. We then look at the relation of these results to other research using a causal Bayes approach to investigate the relationship between causal and conditional reasoning. Finally we look at some of the more theoretical issues raised concerning structure and strength in causal reasoning and in reasoning and argumentation more generally.

Logic and Mental Models Theories

Sloman and Lagnado (2005) have worked through the ability of a variety of theories in the psychology of reasoning to explain the causal discounting effects. These explanations also apply to the common cause condition (CEC) in the present experiment. The common effect condition (CECN) produces augmentation effects because the occurrence of an effect that shares a common cause with other effects raises the probability that they also occur. Augmentation is as inexplicable as discounting from the perspective of logic-based theories.

In the present experiments, participants are asked for probability ratings. Logical probability involves counting the possibilities in which a proposition is true and dividing by the number of possibilities. This account underlies the mental models approach to probability and hence should underpin how mental models explain these data (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). The only alternative is to argue that the models are annotated with probabilities, which places the processes responsible outside the explanatory purview of mental models theory. We now show in detail why discounting cannot be predicted by mental models theory.

The eight logical possibilities for the ECC and CEC cases are given below in (A) and (B), respectively. Those that must be false assuming the truth of the background premises, B, are listed in bold. The premises additionally ruled out by the assertion of the consequent, C or E, respectively, are shown in italic. Logically, no other information is available and so the remaining possibilities must be equiprobable, yielding predictions to compare with our Bayesian predictions and with the data.

Following Byrne (1989) and Byrne et al (1999) we encode the background premises in the alternative cause CEC case as a disjunctive antecedent, i.e., if C1 or C2, then E. We also allow that although the diagnostic ECC conditionals place the cause in the consequent and the common effects in the antecedent clauses, participants may convert the clauses of these conditionals and represent them as a single conditional with a conjunctive consequent, i.e., if C, then E1 and E2. Such conversion is a familiar explanatory manoeuvre in the psychology of reasoning, but of course logically it is an error. Conversion is also consistent with Goldvarg and Johnson-Laird’s (2001) account of causal reasoning in which left to right order of a model reflects causal order, not antecedent-consequent order. Accordingly, with conversion, seven and six of the eight possibilities are ruled out in (A) and (B), respectively.
In (A) those possibilities ruled out by the truth of the conditional, if \( C \), then \( E \)
and \( E_2 \), are in bold only; those possibilities further ruled out by the truth of \( C \) are shown
in bold and italics. As we can see for ECC in (A) only one possibility is left, so \( R \) should
be high, with \( P(E|B, C) = 1 \). This is because this is simply a logically valid MP
inference. Moreover, being then told that \( E_2 \) is the case, cannot change this probability as
it is only true in the same possibility, i.e., for \( R_2, P(E|B, C, E_2) = 1 \). So for ECC, \( \Delta R \)
should be zero and the ratings should be high, which is consistent with the data, \( P(E|B, C) = P(E|B, C, E_2) = .80 \).

In (B), the possibilities ruled out by the truth of the conditional, if \( C_1 \) or \( C_2 \), then \( E \),
are in bold only; those possibilities further ruled out by the truth of \( E \) are in bold and
italic; finally those possibilities additionally ruled out by the truth of \( C_2 \) are in bold and
are underlined. From (B) for the CEC case, it is clear that \( P(C|B, E) = P(C|B, E, C_2) \)
for CEC, \( \Delta R \) should be zero and the ratings should be mid-range. However,
consistent with our Bayesian analysis \( P(C|B, E) = .67 > P(C|B, E, C_2) = .49 \).

For ECNC and CENC, we simply make the same calculations but without
assuming the truth of the consequent of the original conditionals. For ECNC, in (A),
the four possibilities in which \( C \) is false (\( \neg C \)) are also now possible true instances of
the conditional, i.e., \( B \), making 5 in all and \( E_1 \) is true in three of these, so \( P(E_1|B) = .6 \). Once
\( E_2 \) can be assumed to be true, then only 3 possibilities remain and \( E_1 \) is true in two of
these and so \( P(E_1|B, E_2) = .67 \). Consequently \( \Delta R \) should be marginally greater than 0,
which is consistent with the data: \( P(E_1|B) = .59 < P(E_1|B, E_2) = .73 \).

For CENC, by similar reasoning, \( P(C'|B) = .4 \) and \( P(C'|B, C_2) = .5 \), i.e., \( \Delta R \)
should be marginally greater than zero. This prediction is in the opposite direction to the
data: \( P(C'|B) = .58 > P(C'|B, C_2) = .44 \). In sum, this plausible mental models approach
does not make the right predictions.

One might argue that without conversion mental models would fare better.
However, without conversion the CE and EC cases would have to be treated identically
as they are logically equivalent. Consequently, mental models would have to make the
same predictions for both causal orders, which is not consistent with the highly
significant main effect of causal order.

The idea that regardless of the order in the premises, people build a mental
mechanism (Chater & Oaksford, 2006) consistent with the causal order, provides some
motivation for why people may convert these premises. The converted conditional
provides the most natural description of the directed mental representation of the
premises \( B \). Oaksford and Chater (this volume) also discuss how directed or ordered
representations may emerge from interrogating a mental mechanism like a causal Bayes net. However, they used a constraint satisfaction neural network to implement the relations that conditionals may describe. Moreover, they explicitly include connections that build in knowledge of correlated causes (although not of correlated effects). This implementation raises the issue of how far a rational, computational level model can go in explaining causal conditional reasoning. This is because correlated causes violate the independence assumptions of causal Bayes nets (Pearl, 2000).

Of course, it has been a totally standard explanatory tool in cognitive science to appeal to the algorithmic or performance level to explain discrepancies with the computational/competence/rational level. This is the strategy pursued by Oaksford and Chater (this volume). The hypothesis we have explored here is that in human conditional reasoning, people build a mental mechanism (Chater & Oaksford, 2006) that is a dynamic representation of the dependencies described by a conditional which can actively interrogate (see also Oaksford & Chater, this volume).

The Relationship between Causal and Conditional Reasoning

Besides our causal Bayes approach, there is also a related approach by Sloman and Lagrado (2005). These authors, however, argued that conditional reasoning differs in quality from causal reasoning, based on evidence showing that causal descriptions and corresponding conditional descriptions produced various systematic differences. In particular, conditional statements tended to produce weaker effects than causal statements. However, we believe there are several issues with Sloman and Lagrado’s (2005) experiments, to do with an important element of Causal Bayes models, the distinction between structure and strength (Griffiths & Tenenbaum, 2005).

We believe that conditionals, “if C then E,” are structure building operators in the same way as “C causes E,” i.e., they suggest that some dependency exists between C and E. Different information about enablers, common causes or effects, will alter the local mental structure people build in working memory (Oaksford & Chater, this volume). However, to perform further inferences, the various parameters of these structures, or mental mechanisms (Chater & Oaksford, 2006), must be set from linguistic or environmental cues or from prior knowledge. The mental structures built will be the same for causal and other dependencies, reflecting the idea in situation semantics (Barwise & Perry, 1983) that causal dependencies are regarded as the core meaning of the conditional. However, the parameters of these models, their strength, and whether additional machinery, perhaps to do with utilities, are needed will vary between different types of conditional sentences.

We argue that the differences between causal and conditional reasoning that Sloman and Lagrado (2005) observe are due to variations in the strength but not structure. We look at two cases. In their Experiment 1, a structure $A \rightarrow B \rightarrow C$ is introduced concerning three billiard balls. This structure is described using causal terminology, e.g. Ball 1’s (A) movement causes Ball 2 (B) to move, causal conditionals, e.g., If Ball 1 moves, then Ball 2 moves, or logical conditionals, which use the same conditionals but with the preamble, “Someone is showing off her logical abilities. She is moving balls without breaking the following rules.” Participants are then told that there is an intervention on B which prevents it from moving, this should lead to the structure A
B → C. So if asked, “Imagine that Ball 2 could not move, would Ball 1 still move?” participants should say “Yes,” but if asked, “Imagine that Ball 2 could not move, would Ball 3 still move?” participants should say “No.” For all three descriptions participants said that Ball 3 would not move but for the “logical” conditionals about 45% said that Ball 1 would still move whereas 90% endorsed this statement for causal and causal conditional descriptions.

The only difference in the logical condition relates to the possible causes of Ball 1 moving. In the “causal” cases, a causally open system is described where the normal causes of Ball 1 moving are in operation which should be unaffected by Ball 2 being prevented from moving. However, for the “logical” case a closed system is described in which the only cause of balls moving and so of Ball 1 moving is an intentional action of the person showing off her logical abilities. With no knowledge of the rule governing whether she moves Ball 1, participants assume she moves it at random, i.e., the probability is approximately .5. So this is not an instance of people adopting a fundamentally different interpretation of conditional and causal statements. The “logical” preamble simply changes the interpretation of the initiating causes of Ball 1 moving, from an open system to a closed system concerning the intentional actions of an agent.

Differences between causal and conditional reasoning in their Experiment 3 (Sloman & Lagnado, 2005, p. 20) can also be explained by the difference between strength and structure:

“Germany’s undue aggression has caused France to declare war. Germany’s undue aggression has caused England to declare war. France’s declaration causes Germany to declare war. England’s declaration causes Germany to declare war. And so, Germany declares war.

(1) If England had not declared war, would Germany have declared war?
(2) If England had not declared war, would Germany have been aggressive?”

Thus the causal statements categorically assert that Germany’s undue aggression was the cause of France going to war. They are in the past tense describing events that have already occurred. This is in contrast with the conditional statements, using materials like: “if Germany is unduly aggressive, then France will declare war.” The consequent of the conditional statement is in the future tense describing an event that may happen in the future given the antecedent event occurs. The future is uncertain in a way that the past is not. Consequently these descriptions differ in strength but relate to the same causal structure. To compare like with like would involve using causal statements like “Germany’s undue aggression will cause France to declare war.” The difference in strength accounts fully for the lower endorsement of questions 1 and 2 for the conditional case. In sum, Sloman and Lagnado’s (2005) data are entirely consistent with the present Causal Bayes approach to both conditional and causal reasoning.

Conclusions

This chapter has emphasized that patterns of verbal reasoning may, in some cases, be understood in terms of causal, rather than logical, structure. Moreover, we have focused on a case in which the broad pattern of reasoning data observed is predicted by the normative theory of causal reasoning (Pearl, 2000). This approach may apply more generally: indeed, there has been a recent surge of interest in relating normative theories
of causal reasoning to a wide range of cognitive phenomena, ranging from reasoning to learning, and aiming to capture the behaviour of participants including adults (Glymour 2001; Sloman, 2005; Tenenbaum & Griffiths, 2006) children (Gopnik et al., 2001, 2004) and rats (Deow, Courville & Dayan, 2008; Waldmann, Cheng, Hagemeyer & Blaisdell, 2008). To pick a particularly elegant example, Blaisdell, Sawa, Leising & Waldmann (2006) trained rats, in separate trials, to associate a light with either food or a tone. On hearing the tone they expected food; but when they pressed a bar to cause the tone, the tone elicited reduced expectation. Blaisdell et al.’s interpretation was that the rats had inferred that the light was a common cause of the food and tone, and hence that hearing the tone provided evidence that the food might also be present. But when the rat judged its own behaviour to have caused the observation of the food, this inference is blocked.

There is, though, also considerable interest in apparent departures of rational causal reasoning. For example, Quattrone and Tversky (1984) asked people to see how long they could hold their arms in cold water, with the cover story that tolerance was associated either with a healthy, or an unhealthy, heart. Those in the former condition tolerated the cold water for longer, suggesting that people tended to behave in a way that signals good, rather than bad, news about their health. But, of course, by manipulating a putative effect of the condition of their heart, they could not reasonably hope to influence the cause—i.e., the actual condition of their heart. Such apparent irrationality may be less stark, though, if we allow that people take account of the negative psychological impact of receiving “bad news.” Thus, rather than test my endurance to the limit, possibly signalling bad news for my future health, I may not unreasonably choose to drop out early. The extent to which this and other apparent departures from normative standards of causal reasoning can be explained in rational terms, by taking a wider perspective on the reasoning problem, that also includes the issue of reasoning in children and animals, is an interesting question for future research (cf. Hilton, 1990; McKenzie & Nelson, 2003).
References


Table 1: Problem structure and predictions for four experimental conditions (with concrete examples in light gray)

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
<th>Condition 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 -- Cause to Effect Directive; Consequent Not Asserted</td>
<td>C1 -- Cause to Effect Directive; Consequent Not Asserted</td>
<td>C1 -- Cause to Effect Directive; Consequent Asserted</td>
<td>E1 -- Effect to Cause Direction; Consequent Asserted</td>
</tr>
<tr>
<td>Causal structure</td>
<td>C -- Causal</td>
<td>E -- Diagnostic</td>
<td>C -- Causal</td>
</tr>
<tr>
<td>Background information given to participants, B</td>
<td>If C1 then E (Promise 1) If C2 then E (Promise 2)</td>
<td>If E1 then C (Promise 1) If E2 then C (Promise 2)</td>
<td>If C1 then E (Promise 1) If C2 then E (Promise 2)</td>
</tr>
<tr>
<td>If it rains, then the street is wet. If the sprinklers are on, then the street is wet.</td>
<td>If it is warm outside, then it is sunny. If there are shadows, then it is sunny.</td>
<td>If it rains, then the street is wet. If the sprinklers are on, then the street is wet. The street is wet.</td>
<td>If it is warm outside, then it is sunny. If there are shadows, then it is sunny. It is sunny.</td>
</tr>
<tr>
<td>Rating 1</td>
<td>How likely is C1? (Corresponding to Prob(C1, B))</td>
<td>How likely is C1? (Corresponding to Prob(E1, B))</td>
<td>How likely is C1? (Corresponding to Prob(C1, C, B))</td>
</tr>
<tr>
<td>Information given to participants</td>
<td>How likely do you think it is that it rains?</td>
<td>How likely do you think it is that it is warm outside?</td>
<td>How likely do you think it is that it rains?</td>
</tr>
<tr>
<td>C2 is the case (Promise 3)</td>
<td>E2 is the case (Promise 3)</td>
<td>C2 is the case (Promise 4)</td>
<td>E2 is the case (Promise 4)</td>
</tr>
<tr>
<td>The sprinklers are on.</td>
<td>There are shadows.</td>
<td>The sprinklers are on.</td>
<td>There are shadows.</td>
</tr>
<tr>
<td>Rating 2</td>
<td>How likely is C1? (Corresponding to Prob(C1/C2, R))</td>
<td>How likely is C1? (Corresponding to Prob(E1/E2, R))</td>
<td>How likely is C1? (Corresponding to Prob(C1/C2/E, R))</td>
</tr>
<tr>
<td>How likely do you think it is that it rains?</td>
<td>How likely do you think it is that it is warm outside?</td>
<td>How likely do you think it is that it rains?</td>
<td>How likely do you think it is that it is warm outside?</td>
</tr>
<tr>
<td>Prediction after Pearl (2000)</td>
<td>( P(C1/C2) = P(C1) ) Independence</td>
<td>( P(E1/E2) &gt; P(E1) ) Augmenting</td>
<td>( P(C1</td>
</tr>
</tbody>
</table>
Table 2 Diagrams showing the causal relationships between the variables within each scenario.

<table>
<thead>
<tr>
<th>Causal Direction</th>
<th>Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flower</td>
</tr>
<tr>
<td>Cause-to-effect</td>
<td>Sun → Flower</td>
</tr>
<tr>
<td></td>
<td>Water</td>
</tr>
<tr>
<td>Effect-to-cause</td>
<td>Sun ← Blue Flower</td>
</tr>
<tr>
<td></td>
<td>Red Flower</td>
</tr>
</tbody>
</table>
Figure 1. Predictions showing the main effects for causal direction (cause-effect [CE], effect-cause [EC]) and consequent (Consequent [C], Non-Consequent [NC]) conditions.
Figure 2. Results of Ali et al’s (2008) Experiment 1 showing the main effects for causal direction (cause-effect [CE], effect-cause [EC]) and consequent (Consequent [C], Non-Consequent [NC]) conditions
Figure 3. The button-and-light boxes used to illustrate common effect and common cause structures for children; both structures were implemented in two different scenarios; each child saw the common effect version of one scenario and the common cause version of the other.
Figure 4. Results of the Experiment showing the main effects for causal direction (cause-effect [CE], effect-cause [EC]) and consequent (Consequent [C], Non-Consequent [NC]) conditions.