ACCEPTED VERSION

Kieran C. O'Doherty, Daniel J. Navarro and Shona H. Crabb

A qualitative approach to the study of causal reasoning in natural language: The domain of genes, risks and cancer

Theory & Psychology, 2009; 19(4):475-500

10.1177/0959354309336321

A Qualitative Approach to the Study of Causal Reasoning in Natural Language: The Domain of Genes, Risks and Cancer

Kieran C. O'Doherty (kcodoher@interchange.ubc.ca)
W. Maurice Young Centre for Applied Ethics, University of British Columbia
227 – 6356 Agricultural Road, Vancouver, BC, V6T 1Z2, Canada

Daniel J. Navarro (daniel.navarro@adelaide.edu.au)
School of Psychology, University of Adelaide
Adelaide, SA 5005, Australia

Shona H. Crabb (shona.crabb@adelaide.edu.au)
Discipline of Psychiatry, School of Medicine, University of Adelaide,
Adelaide, SA 5005, Australia

Abstract

Causal reasoning has been studied extensively in experimental cognitive psychology. Generally, the focus is on how individuals learn causal relationships in their environment through observation or interventions. Although it seems self-evident that causal beliefs about some phenomena are learnt largely through linguistic channels, to our knowledge no empirical studies have addressed this issue. In this paper we investigate causal reasoning that is embedded in naturally occurring language. We focus on genetic counselling for cancer, in which complex relationships between genes, medical interventions and cancer are communicated by health professionals to clients. We borrow the idea of graphical causal maps from previous experimental studies and show that they can be applied to the study of causal reasoning in naturally occurring talk. We see this study as complementing existing experimental research, while maintaining that the study of causal structures embedded in naturalistic language adds an important dimension to our understanding of causal reasoning.

Key words: Bayes nets, causal reasoning, causal maps, discourse analysis, risk, genes, cancer, genetic counselling, learning

Biographies

Kieran O'Doherty is a Research Associate at the W. Maurice Young Centre for Applied Ethics, University of British Columbia. Research interests include discourse analysis, agency and decision-making, the management of risk and uncertainty, and the use of public engagement and deliberative democracy as a methods for securing public input into policy. ADDRESS: W. Maurice Young Centre for Applied Ethics, University of British Columbia, 227 – 6356 Agricultural Road, Vancouver, BC, V6T 1Z2, Canada. [email: kcodoher@interchange.ubc.ca]

Daniel Navarro is an Australian Research Fellow at the University of Adelaide. His research examines the mathematical structure of human conceptual knowledge and decision-making, and develops statistical methods for psychological data analysis. In addition to pursuing basic research into these areas, he maintains an interest in dealing with applied problems across a range of areas that include defence, telecommunications and the petroleum industry. ADDRESS: School of Psychology, The University of Adelaide SA 5005, AUSTRALIA. [email: daniel.navarro@adelaide.edu.au]

Shona Crabb is a Senior Research Officer in the Discipline of Psychiatry, School of Medicine at the University of Adelaide. Her research interests are in qualitative and critical health psychology, with recent projects examining discursive constructions of risk and responsibility in accounts of preventative health practices, and the construction and communication of genetic information. ADDRESS: Discipline of Psychiatry, The University of Adelaide SA 5005, AUSTRALIA. [email: shona.crabb@adelaide.edu.au]

Acknowledgements

KCO was supported by ARC grant LP-0562206 during part of this project, jointly funded by the Australian Research Council and m.Net Corporation. DJN was supported by an Australian Research Fellowship (ARC grant DP-0773794). We would like to thank Graeme Suthers, Joya McCormack, Jacquie Armstrong, Debra Trott, and Sally Russell for indispensable assistance in data collection and much more. We would also like to express our gratitude to all the clients of the Familial Cancer Unit who consented to participate in this study. Finally, we would like to acknowledge the valuable comments of two anonymous reviewers.

A Qualitative Approach to the Study of Causal Reasoning in Natural Language: The Domain of Genes, Risks and Cancer

Causal models provide a natural method for understanding the dependencies that exist between variables in the world (Pearl, 2000), so it is not surprising that human learners tend to form beliefs about the causal relations between entities (e.g., Sloman, 2005). Indeed, while understanding how people uncover causal laws is a topic of study in its own right in psychology (e.g., Cheng, 1997; Gopnik et al., 2004; Griffiths & Tenenbaum, 2005; Sloman & Lagnado, 2005; Steyvers, Tenenbaum, Wagenmakers & Blum, 2003), causal relations play an important role in other areas, such as the study of conceptual structure (e.g., Rehder, 2003) and conditional reasoning (e.g., Evans & Over, 2004). When modelling this process, it is typical to assume (generally for the sake of simplicity) that mental representations of causal laws develop as a consequence of direct observation of the environment. In everyday life, however, it is very common for people to acquire causal representations through reading or conversation, rather than via direct experience of the environment. To the extent that linguistic data provide proxy 'observations' for people, this is unlikely to pose any particular difficulty for psychological models, although it certainly complicates matters. On the other hand, the ready availability of linguistic data provides a rich source of data regarding the naturalistic use of causal concepts. In this paper we build on previous work on the cognitive processes assumed to underlie causal reasoning. In particular, we draw on the notion of causal maps as useful representations associated with causal reasoning processes (Gopnik et al., 2004). However, rather than provide participants with a simple environment in an experimental setting in which we can control the causal relationships and attempt to model behaviour, we take a more exploratory approach and focus on some of the ways in which people employ causal language in spontaneous real-world speech. The advantage of this approach is that we observe causal language in a natural environment, but the disadvantage is that we forgo control and quantification.

The domain in which we undertake our investigation is genetic counselling sessions for familial cancer. We focus on the causal representations that are implied by the genetic counsellors as these may be considered to be the 'stimuli' upon which clients' subsequent causal representations are likely to be based. In order to analyse these data, we employ qualitative methods commonly used in social psychology and the health sciences (e.g., Wetherell, Taylor & Yates, 2001). In line with common applications of discourse analysis, our commentary is focused primarily in the domain of interactive communication, rather than individuals' actual cognitive processes. Rather than make the strong assumption that causal models are literally communicated in speech, we make the weaker suggestion that some causal representations may be viewed as being 'implicit' in speech, and act to constrain the interpretation of events that a listener might entertain.

As we are interested in a knowledge domain (familial cancer) in which causal relations are transmitted to a lay population almost exclusively via discourse, we feel that an investigation that focuses on this discursive dimension is not only justified, but essential. We see this study as complementing experimental studies that investigate causal reasoning processes in a controlled environment. Whereas the value of experimental studies lies primarily in understanding the formation of causal representations resulting from direct interactions between individuals and their environment (e.g., Steyvers, et al. 2003), the present study contributes to an understanding of the causal frameworks formed as a result of communication with others. We also see this study as adding to experimental studies by investigating the operation of phenomena that have been studied extensively in laboratory contexts in more naturalistic settings (e.g., Neisser, 1981).

Representing Causal Relations

Recent views of causal models in cognitive psychology have borrowed heavily from the graphical model formalism advocated by Pearl (2000). While it is neither necessary nor desirable to outline this approach in any mathematical detail, a brief qualitative overview may be useful. The simplest form of causal model takes the form of a *directed acyclic graph* that connects variables, first suggested by Wright (1934), and illustrated in Figure 1. This network is derived by considering the following dialogue from *The Simpsons*:

Homer. Every time I learn something new, it pushes some old stuff out of my brain. Remember when I took that home wine-making course, and I forgot how to drive?
Marge. You were drunk.

---- Figure 1 about here ----

The cleanest representation of this interaction involves four variables, corresponding to the act of taking the winemaking course, learning new facts, getting drunk, and forgetting how to drive. Homer's statement sets up the following causal relations: (1) wine-making education causes new learning, and (2) new learning causes forgetting of prior learning. Marge's reply suggests that, instead, (3) the wine-making knowledge causes drunkenness, and (4) drunkenness impairs driving. Notice firstly the *directed* nature of the connections. Impaired driving is an effect of drunkenness, not its cause. So improving Homer's driving skills will not cause him to sober up, but sobering up *will* improve his driving. This asymmetry lies at the heart of causal explanation. Secondly, note that the causal relations are acyclic: there is no path by which the winemaking course can indirectly cause the winemaking course. Finally, note that the two potential paths by which winemaking courses can cause poor driving (skill loss and alcohol) are competitive. By postulating an intoxication-based explanation for Homer's driving,

Marge is implicitly refuting Homer's "limited capacity" explanation. This phenomenon is called *explaining away*, and plays an important role in understanding human causal reasoning.

Gopnik et al. (2004) refer to these graphs as *causal maps* and argue that they are an improvement over other explanations for learning causal relationships in the world as they allow for people to learn causal relations by observation, rather than relying solely on learning by intervention, in a trial-and-error fashion (on this topic, see also Lagnado & Sloman, 2004; Lagnado, Waldmann, Hagmayer, & Sloman, 2006). We would like to go one step further and suggest that in many instances, causal representations are formed without direct learning or observation. In our view, causal representations are also formed through participating in communicative activities, or being the target of certain communications.

In the present study, we use causal maps in much the same way as Gopnik et al (2004). As Gopnik et al point out, although causal maps may represent causal knowledge, they are not necessarily the only devices that perform such a function. Nevertheless, they are a likely candidate when it comes to learnt relationships. Importantly, while causal maps may encode deterministic relationships between variables, in general they encode probabilistic relationships. One significant difference between the processes for learning causal maps considered by Gopnik et al. and our use of causal maps here is that Gopnik et al. are concerned with inferring causal maps from observed events (the inverse causal problem). Clearly, this is both a conceptual and a computational problem. In contrast, we are here concerned with learning causal maps not from discrete data, but from discourse in which *complete causal maps* are already implicitly embedded.

Although we propose a genuinely novel approach to this problem, there has been some relevant work on the study of causal reasoning through language, both in cognitive psychology and other disciplines. Drawing on literature from cultural and medical anthropology, for instance, Lynch and Medin (2006) study the ways in which different groups construct causal chains about illnesses. The authors convincingly argue that, "illness explanatory frameworks are not necessarily tied to single cognitive domains and that the notion of cognitive domains is not sufficient to explain how people construct causal models of illness." (p.306). Significantly, Lynch and Medin's study drew on interview data, in which participants were asked explicitly to draw causal relations between factors. Other recent work in cognitive psychology has focused on learning of categories and causal relations through language. For example, in one study Sloman, Love, and Ahn (1998) asked participants to draw links between concepts with which they were already familiar. Similarly, much of the literature on conditional reasoning (e.g., Evans & Over, 2004) considers the implied role of causality in language. For instance, in the study of counterfactual conditionals (e.g., Over, Hadjichristidis, Evans, & Sloman, 2007), the role played by causality is central: while a typical conditional of the "if P then Q" kind often implies a causal relation in a linguistic form, the counterfactual version "if P had happened then Q would have happened" makes the causality somewhat more obvious. Finally, some recent studies have focussed on how causal

relationships are expressed linguistically (e.g., Wolff, 2003). Most significantly, Wolff, Song, & Driscoll (2002) conducted an important study that focussed on how models of causation capture distinctions that people make when using causal verbs, such as "hinder", "help", and "prevent".

While we do not wish to detract from the valuable insights gained in these studies, we would like to emphasise some important differences to our approach. First, while Lynch and Medin's (2006) data are also qualitative in nature, our study is quite different in that we focus on the process whereby causal beliefs are transmitted via language. We thus focus on the causal structures that are embedded in natural language, rather than eliciting them explicitly. As we demonstrate below, this has important consequences, not least of which is that we find the language about causal relationships in our domain to be highly ambiguous (something that would not have emerged, had participants been asked to state their beliefs about causal relations explicitly). Second, because of the experimental design of such studies as Over et al (2007), these statements are examined in isolation and under controlled conditions. In our study, we forgo the tight control offered by traditional experimental designs in favour of a 'field study' that offers a real world context (and we hope to demonstrate that this trade-off is worthwhile, given some of the patterns we are able to demonstrate through our analysis). We should also point out here that we use the term 'natural language' somewhat differently to the way it is employed by Over et al (2007) and authors of other similar studies. In particular, Over et al state that they use 'natural language' in their study of causal reasoning. However, while their study uses statements that are 'natural' in the sense that they are meaningful expressions written in standard English rather than, say, propositional logic, this is still very different from the kind of real-life environments in which we are interested. We are interested in natural language in the sense that it is naturally occurring. In fact, as our data consist of audio recording of genetic counselling sessions, they were not even elicited for the purposes of the study, but constitute a snap shot of how causal structures are embedded in and transmitted through language in real world contexts. Finally, the analysis we present below offers a level of detail that is not available in the application of software packages to large data sets.

In summary, Gopnik et al. (2004) argue that people, and in particular children, represent causal relationships in ways that can be described by causal Bayes nets. They also argue that individuals construct new causal representations through observing correlations and interventions. To a large extent, therefore, Gopnik et al. are concerned with the process whereby individuals come to acquire beliefs about causal relationships between certain phenomena**.** In contrast, we are not concerned with the question of how individuals come to construct causal representations from first principles. We agree with Gopnik et al.'s notion of causal maps, but argue that in many instances causal maps need not be constructed through observations of correlations, and interventions. In many cases, already constructed causal representations are conveyed to individuals through some process of communication. In the remainder of this paper, we focus on the causal maps that are embedded in one such process of communication.

Causality and Familial Cancer

In familial cancer clinics, individuals and families attend genetic counselling sessions to discuss the possibility that they may have a genetic predisposition towards specific forms of cancer. Based on these consultations, individuals and families have to make decisions about taking genetic tests, implementing risk management strategies, medical surveillance options, and disseminating risk information to other relatives (O'Doherty & Suthers, 2007). The domain is psychologically interesting for a number of reasons. Firstly, there is a great deal of uncertainty as to the 'correct' causal model for the development of cancer in any given individual. This uncertainty is reflected in the talk of counsellors and clients, which can often be seen to reflect a range of hypothesized causal models. Secondly, while both clients and counsellors possess a rich domain theory for cancers and related phenomena, not all of the variables of interest have well-defined roles in the domain. For instance, while *genes*, *cancers* and *surgery* are objects that fit easily into people's existing knowledge via physical domain theories (e.g., Griffiths, Baraff & Tenenbaum, 2004), *risks* do not have a straightforward interpretation (O'Doherty, 2006). The ambiguity associated with the concept of risk introduces interesting variation in causal talk. As a result, in different parts of the conversation, a range of different causal structures are implied. Thirdly, while most of the speakers have some experience of cancer, either through personal experience or via close relatives, the causal factors behind cancer are sufficiently complex that most of the *causal* knowledge people possess must be inferred through linguistic interactions. These features contribute to the suitability of the domain of genetic counselling as a basis for an investigation into the latent causal structures employed in language.

Our data consist of transcripts from a number of genetic counselling sessions provided by the Familial Cancer Unit of a large Australian hospital. The complete corpus contains over 30 sessions. Extracts are presented in 'semi-sanitised' form. That is, while all utterances are shown as they appear in the recording, fine grain details such as pauses and changes in pitch are omitted; punctuation has been added to enhance readability. All names appearing in the extracts are pseudonyms; however, gendered pronouns accurately indicate speakers' gender. Parts of the transcript are underlined to draw attention to particular features.

Our approach is to present a number of different extracts from the corpus, each of which suggests one or more implicit causal maps of the phenomena under consideration. Our analysis focuses on examining the range of causal maps that are consistent with the information presented in the extract. Of course, there is an interpretive element to this analysis which is unavoidable given the nature of our approach. Accordingly, caution is required, particularly when considering the extent to which a listener would actually "receive" a causal model from the speaker. Nevertheless, while the literal transmission of a causal model seems highly unlikely, we suggest that the language used by a speaker acts to constrain the set of plausible causal models and, therefore, the range of interpretations available to a listener.

Linking Events, Risks and People

Individual differences pose a complex statistical problem for psychological researchers. No two people are ever likely to possess precisely the same competences for a particular task, so any data they furnish are unlikely to be drawn from the same statistical distribution. One consequence of this issue, well-known in the literature (e.g., Estes, 1956), is that averaging across observations of different people can sometimes produce data that are unrepresentative of any individual. Yet, without some way of aggregating information across people, it becomes impossible to draw any reasonable conclusions about data. While there are a number of sophisticated solutions (e.g., Rouder, Sun, Speckman, Lu & Zhou, 2003; Navarro, Griffiths, Steyvers & Lee, 2006), the issue can hardly be said to be resolved.

The same tension can be observed in the causal language used in genetic counselling sessions. The prior knowledge available to speakers (or at least to counsellors) is epidemiological in nature and aggregates data from a large number of people. However, this information is useful to the client only insofar as it relates to the question of whether they personally will develop cancer. Naturally, the extent to which these data generalize to the individual depends heavily on the kinds of causal processes that actually govern cancer. However, there is a great deal of uncertainty about these processes and, in all likelihood, the correct causal model is exceedingly complex. Given this, it is interesting to examine the causal assumptions that are implicit in the spontaneous speech of clients and counsellors.

A good example is provided by the following extract. The client has sought genetic counselling due to an unknown condition for which she had already undergone surgery. The passage occurs near the end of a session and follows a discussion in which the geneticist indicated to the client that, although he is not able to give her any more knowledge regarding the nature of her condition, he is fairly certain that it is not cancer related. Present in this session were the geneticist, a genetic counsellor ('Jane'), the client and the client's husband.

Extract 1

Geneticist. Now, just to warn you that (inaudible) for me and Jane (the genetic counsellor) and others, the risk of developing cancer is of the order of 1 in 3. That's the standard risk you run, I run, we all run. I'm suggesting that your risk of getting cancer is about 1 in 3 as well. So, it's not zero but it's not necessarily higher than for anyone else. And at this stage I do not think that we need to do anything special for your children or your brothers or sisters (inaudible) special tests or whatever on them, eh, because I don't at this stage have evidence of something that places them at risk of problems.

Extract 1 illustrates a prototypical explanation given to clients by geneticists and genetic counsellors for cases in which there are no indications of a genetic predisposition to cancer. Consistent with Gricean maxims (Grice, 1975), the talk is as minimal as required to convey the 'appropriate' intuition, and no explicit reference is made to any actual causes of cancer. The only data seen to be relevant are those that govern base rate or

'background' risk. However, there is some ambiguity in the way in which these base rates operate. Initially, the geneticist refers to the 'standard' or baseline risk, which is represented as affecting most people: "*me and Jane and others … the standard risk you run, I run, we all run"*. The universality and invariance implied by this use of three-part lists (Jefferson, 1990) suggests that this statement should be interpreted in a manner consistent with the *shared risk* model shown on the left of Figure 2a. In this case, we postulate a single fixed risk, corresponding to the universal probability of developing cancer. Since the risk is shared by all people, the epidemiological data are assumed to be directly relevant to the estimation of the risk facing the client ("*your risk of getting cancer is about 1 in 3 as well"*).

---- Figure 2 about here ----

In contrast, consider the phrasing that follows immediately after, suggesting that the probability that the client will develop cancer is "*not zero, but it's not necessarily higher than for anyone else"*. This statement is not consistent with the shared risk model, which does not allow any variation in risks across people. If the speaker genuinely intended to convey an intuition that accorded with a shared risk model, this statement conveys no information. For the statement to serve a pragmatic purpose, the "implicit model" under consideration must allow for the possibility of individual differences in risk, a feature not present in the shared risk model. In order to do so, the causal map needs to be expanded, to cover the *i.i.d. risk* scenario illustrated in Figure 2b. In this model, people do not share risks: instead, everyone has their own unique cancer probability. However, without any other information specified, the counsellor suggests that the client faces a risk that is drawn from the same distribution as the rest of the population. In other words, the risks are *independent and identically distributed* (*i.i.d.*).

The shift from shared risk to identically distributed risks is seamless in everyday language, but highly significant in statistical terms. Firstly, it requires a shift in the notion of probability (O'Doherty, 2006; 2007). In the shared risk case it is simple to interpret the latent probability θ as a long-run frequency, as advocated by a number of authors (e.g., Von Mises, 1957). Once individual differences are introduced, this interpretation is unsustainable, since replications of once-off events are no longer possible. Rather, it now seems most natural to think of this probability as a state of uncertain knowledge (e.g., Jaynes, 2003). Indeed, after shifting to the *i.i.d* model, it is implied in this extract that 'evidence' can contribute to an individual having a different (increased) risk: increased or more certain knowledge can change the probability of cancer. This feature can be seen particularly in the use of the phrase "*at this stage"* which is repeated twice in the extract: the risk for the client (and their family) is represented as being reflective of the baseline risk until some 'evidence' changes their risk (*"I don't at this stage have evidence of something that places them at risk"*). This notion of cumulative knowledge changing risks – the epistemic nature of risk – appears frequently in genetic counselling sessions and is certainly linked with the causal models implicit in the talk.

Secondly, the shift from shared to *i.i.d.* risk models is significant because the 'link' between people is now somewhat more abstract. Rather than having a shared risk θ that links a collection of outcomes x, we have some (unarticulated) background factors that link a collection of individual risks θ, each associated with only a single outcome. Under this elaborated view, the counsellor indicates that 1 in 3 is a kind of average risk, which will be expected to vary from case to case.

Curiously, the intuitive concept of risk appears to cover this shift with little difficulty. Thus, when considering this transition, it should be noted that the two accounts shown in Figure 2 are not really competitors. If the true causal structure for cancer is as complex as we suspect it is, everyday reasoning must rely on useful approximations to this (unknown) structure. The shift from a shared risk account to an *i.i.d.* risk account should thus be interpreted not as a change in belief about the true nature of cancer but as a shift in the level of description relevant to the discussion. This can be seen by considering the final sentence in the extract, in which the counsellor refers to the possibility of correlated risks faced by client's siblings. Implicit in this shift is the recognition that risks are not *independently* distributed when there is covariation in background causal factors (e.g., genes). It is to this topic that we now turn.

Linking Genes, Risks and Cancer

In the analysis of the first extract, our aim was to illustrate the implied relationships between individual risk and population data, shown most clearly in a case where no apparent risk factors exist. However, in the context of genetic counselling it is common for explanations to invoke the genetic inheritance of risk. By doing so, speakers are able to elaborate on the models discussed earlier. By introducing genetic factors to the causal models, it becomes possible to allow structured relationships between people to emerge in the talk: shared genes lead to similar risks. Although this level of explanation is implicit in Extract 1 (in the reference to the possibility of suggesting 'special tests' for brothers and sisters), the following two extracts illustrate this more clearly.

Extract 2 comes from a session in which a client seeks advice about her chances of developing breast cancer. The extract is taken from early on in the session and follows a discussion about the process of medical surveillance (which would become relevant should the client be diagnosed as being at increased risk of breast cancer). Extract 3 comes from a session involving a genetic counsellor, a client and her husband. The extract follows a discussion about the implications should the client test positive for a genetic mutation known to predispose the carrier to developing cancer.

Extract 2

Counsellor. We wouldn't know if you're going to get breast cancer in your lifetime.
We know that this gene might place you at increased risk but we don't have any magic way of working out whether you're going to get breast cancer, though.

Extract 3

Counsellor. And so in a short period of time information changes and that will be the same for the future. We know this much today but in a few years time we'll know even more. So it is dependent on the day, to a degree. If there is things that we learn in the future that are important for you we always try and keep in touch. So, if you move again always let us know. Put us on your regular mailing list for change of address so that we can keep in touch down the track. tsk. If we find an abnormality in that gene it would mean that there is a risk further problems happening to you and that we can get a bit more specific about that whereas at the moment I haven't told you anything. And that's why it would be important that we see you again.

When genes are used in causal explanations for the potential development of cancer there are at least three alternative mechanisms that can be attributed to them, illustrated in Figure 3. We refer to the model on the left as the *strictly causal* explanation, in which the individual's genes are treated as an indirect cause of cancer: the presence of a particular genetic factor leads directly to an increased risk. An obvious alternative to this is a *strictly symptomatic* model, in which risks and genes share a common cause. This explanation, illustrated in the middle panel, implies that the observed genetic abnormality is caused by a set of latent factors *f* that also contributes to cause cancer. The third alternative is the *correlational* model illustrated on the right, in which risks are assumed to be correlated with genetic abnormalities, with no causal relationships represented between the two. Note that while these three are by no means exhaustive, they are fairly illustrative of the class of models we wish to consider: we assume that people's domain theories do not allow a cancer outcome to cause a genetic profile, for instance.

---- Figure 3 about here ----

Examining the data presented in Extracts 2 and 3 allows us to consider whether, and in which ways, such causal representations may appear in talk. For example, Extract 2 provides a good example of an explicitly causal statement, through the suggestion that "*this gene might place you at increased risk*". In a purely correlational or common cause context, it would seem strange to use this kind of language. One does not suggest that "*yellowed finger tips might place you at increased risk of lung cancer*" (since both are caused by smoking), or that "*the year in which you took this IQ test might place you at risk of a lower score*" (the "Flynn effect" correlation). It is not that either statement is difficult to understand, but both seem to violate basic pragmatics of everyday English. To be "*placed at risk*" by some factor (e.g., "*this gene*") tends to imply direct causality and agency on the part of the factor. This causal representation in the talk would thus seem to best reflect the strictly causal model in Figure 3a. Contrast this language with the rather more noncommittal remark in Extract 3, in which the presence of "*that gene … would mean that there is a risk of further problems.*" In this context, the speaker does not suggest any noticeable causality, only implication. Similarly, when discussing the Flynn effect, it

seems appropriate to remark that "*the fact that you took the test a long time ago might mean that you scored lower than the current norms would predict*". In this case, the phrase "*would mean that*" carries nothing stronger than the logical implication that might be suggested by a simple statement about conditional probability (Figure 3c).

One of the problems in learning causal structures through verbal communication is that the causal relationships between events and variables are rarely the explicit subjects of communication. Thus, causal structures are more often implicitly embedded in discourse, rather than overtly claimed. There is therefore an inherent ambiguity in many statements containing causal information. This ambiguity is particularly evident in talk about genes and cancer in genetic counselling sessions. Indeed, in both the extracts presented above the claims that could be understood as related to causation are often hedged (e.g., "*this gene might place you at increased risk*"). However, although speakers' causal claims may be ambiguous (or even explicitly limited - "*at the moment I haven't told you anything*"), their talk frames the causal representations that are subsequently available to the listener. Our claim, therefore, is not that particular statements literally encode a causal explanation, but that the pragmatic aspects of the conversation will tend to favour one version over another.

Causal Interventions, Counterfactuals, and Repeated Outcomes

The previous sections deal with some of the basic notions that appear to pervade even 'simple' discussions of genes, risks and cancer. In a number of situations, the conversation becomes much subtler, and invokes some interesting counterfactual reasoning about the effect of medical treatments and the way that events might unfold over time. For example, certain events and interventions may change the probability of future events, while leaving the causal structure of variables unchanged. In other instances the structure of causal connections may be altered by an event or intervention. The discussion in this section considers some of these issues and is based on the exchange reproduced in Extract 4. The extract involves a counsellor and two clients, Karen and Terry, who are sisters. Terry has had breast cancer, which was treated successfully through non-surgical means.

Extract 4

Karen. As Terry has still, you know, she hasn't had the [prophylactic] surgery, and she still has both breasts and that sort of thing, can it reappear in both, obviously she can have that gene in the other breast as well. But can it reappear in the same one?

Counsellor. The inherited, I guess, that is the very real difference between inherited cancer and cancer that just occurs by chance. The chance of it occurring is quite small but once you start looking at these higher risk levels, you're looking at an error, or that first step, being copied into all the cells in your body, 'cause we assume that it's made out of that single cell that you started from, it gets passed

> down through the family, then it's a part of every cell in your body. So, it means that the other breast cells also have that same increased risk for going on to develop into cancer. So, yes. That's where the story of your Aunt Belinda, having breast cancer in two, you know, raises the stakes a bit for your father's side of the family. Coz yes, the other breast cells have that increased risk. Now I should say as well though, if we're talking about this error being copied into all the cells in the body, the cancer risk isn't everywhere. The increased risk isn't everywhere. You know, different genes function in different cells. Other cells have other genes that are expressed in them. So we are talking about quite specific risks. And so yes, that increased risk is there for breast cancer. I guess at this moment in time that's of the smallest concern for, for you at this time. The largest concern is being able to stomp on those cells that are there. And get control of them.

One interesting feature of this extract is that the counsellor appears to find Karen's question very difficult to answer. The reason for this difficulty seems to be that, upon examination, the question turns out to be remarkably complex in terms of implied models of causality. Our discussion will therefore focus primarily on aspects of Karen's question and possible underlying causal models, and will begin by clarifying some features of the question.

Initially, the question seems to be about whether it is possible for Terry, subsequent to treatment for her first cancer, to develop cancer in either of her breasts: "*can it reappear in both*". That is, if we take 'it' to refer to cancer, the question can be read as relating to the possibility of the reoccurrence of cancer, and to the impact of the previous treatment on any possible future development of cancer. However, Karen proceeds (whether accurately or not) to answer her own initial question ("*obviously she can have that gene in the other breast as well*"), and to re-ask a slightly different question: "*can it reappear in the same one?*" Her response to her own question thus incorporates the assumption that Terry may still develop cancer in the breast that has, so far, been unaffected and changes the focus of the question on whether it is possible for cancer to reoccur at a previously affected site.

Given the complexity of Karen's question, the underlying causal representations are also very complicated and need to incorporate a range of interrelated elements. These include: the causal relationship between genes and cancer; whether two distinct locations (i.e., both breasts) can have different risks of the same outcome (cancer); the potential for causal models to change over time with the incorporation of interventions (i.e., the effect of a treatment intervention on the subsequent risk of cancer in either breast, and the potential impact of a future prophylactic intervention); and conceptualizing cancer as a discrete series of potential event as opposed to a single event. We will introduce these various elements piecewise.

The first aspect we will consider is *time*, illustrated in Figure 4. In previous extracts in which the outcome under consideration is simply whether cancer develops (rather than *when* it develops) time has not been made explicitly relevant in the causal maps (see, for example, Figure 3a). If we now incorporate it, time can be accounted for by a simple

causal representation like that in Figure 4a. In this model, the outcome can be different at different times, however the causal structure (including the probability of the outcome occurring) remains unchanged across time. (Note that for the sake of simplicity we show only two points in time, with the variables x_1 and x_2 denoting the presence or absence of cancer at each point.) However, Karen's question suggests that the *risk* of a new cancer in the same location at a different time may have changed over time (not just the outcome). The causal representation is thus expanded as shown in Figure 4b.

---- Figure 4 about here ----

The second elaboration required is *location*. The question requires differentiation between the risks associated with the left breast and the right breast, so the simplest (time-dependent) model for this is the one shown in Figure 5a. In this model, each breast is associated with a single risk, on the assumption that these risks are both shaped by genetic factors. The model thus allows for different outcomes for each breast across time. However, since the question implies the possibility that the risk for each breast can change over time, we need to elaborate on this model, and assume the existence of some unspecified breast-specific factor ϕ. Thus, in Figure 5b, each breast has some unique factor ϕ, which is influenced by genetic factors g, but with a risk of cancer θ, that can change over time. In other words, the expansion of the model from 5a to 5b captures the possibility not only of variation in outcomes over time, but also of variation in (breast-specific) risks over time.

---- Figure 5 about here ----

Although not a focus here, it is worth mentioning that the distinction between genetic factors, the breast-specific factor, and the risk of developing cancer may function to allow a causal representation that makes sense of the concepts of genetic abnormalities, genes, and genetic expression: although a gene exists in every cell, the expression varies as a function of cell type, and a specific abnormality might be highly localized. It seems to be this confusion that leads the counsellor to remark initially that "*the other breast cells have that increased risk*" (due to an inherited allele) but that "*the increased risk isn't everywhere*" (due to differential expression), while holding to the distinction between "*inherited cancer and cancer that just occurs by chance*".

Note that all of the discussion up to this point has served mainly to introduce sufficient conceptual machinery to allow us to make sense of Karen's question. Using the six models shown in Figure 6, we attempt more specifically to illustrate the complex reasoning implicit in the question. The first panel (Figure 6a) aims to provide a rough approximation to the state of the world at the time of Terry's initial diagnosis (assuming the cancer was in the left breast). At this stage, we have cancer in the left breast (x_{11}=1) but not in the right (x_{21}=0). Subsequent to that diagnosis, an intervention was staged in the form of some treatment, denoted T. Karen's ultimate question can be seen as relating to whether it is possible for cancer to reappear in this same location (the left breast)

subsequent to it already having occurred and being treated. In part then, she is asking about the effect of the intervention T on the existing causal structure.

One simple possibility would be that T intervenes at the most local level, removing the cancer but doing nothing else. In this case, the intervention disconnects x_{11} from θ_{11}, and resets x_{11} to 0. The future risk of developing cancer in this breast (or the other) is, therefore, not affected. This possibility is illustrated in Figure 6b, and would probably be the most likely candidate had the previous treatment been very local surgery. However, from all appearances, the treatment was something more abstract (probably chemotherapy or radiotherapy), and it is quite possible that Karen was uncertain about the precise effects of this kind of intervention. For example, a second possibility is that the treatment affected the current state of both breasts (chemotherapy would have just such an effect, though obviously real treatments are *uncertain* interventions), as illustrated in Figure 6c. In this case, T sets both x_{11} and x_{21} to 0, disconnecting the (past) state of both breasts from their usual causes. Again, in this case, the future risk of cancer for either breast is not changed by the intervention.

These two are not the only possibilities at hand. Figures 6b and 6c differ in terms of whether the intervention T is *location-specific*, but both assume that it is *time-specific*, in the sense that it does not affect the probabilities of future cancers in either location. An alternative way of conceptualizing the intervention is to assume that it alters the breast-specific factors and, therefore, the future risk of cancer, which gives us the two models shown in 6d and 6e. The model in Figure 6d assumes that the treatment has a permanent (or maybe just persistent) effect, localized to the left breast. In contrast, the model in Figure 6e assumes that T makes a persistent change to both breasts. Although we know that Terry did not have surgery, a mastectomy (unilateral for Figure 6d, and bilateral for Figure 6e) is an example of an intervention that would affect breast-specific factors and change the future risk of developing cancer. (Please note that we have oversimplified this issue somewhat in Figures 6d and 6e by setting all probabilities to zero; even with prophylactic mastectomies there is still a small chance of breast cancer occurring.)

---- Figure 6 about here ----

At this stage, based on Karen's question, we have four possible means to think about the way in which Terry's previous cancer was successfully treated. Understanding these causal models, and the variations between them, could have important consequences for the ways in which any future preventive act (such as prophylactic surgery) is considered. Importantly though, unless one has a clear understanding of the science underlying the treatment T, it will not be clear which of these four provides the best account of the situation that Terry faces when making the decision about whether to have prophylactic surgery (S). Clearly, if model 6e is correct, then there is no need for the preventative surgery at all, since Terry's risk of future cancer in either breast is effectively removed, and any discussion of such surgery would be highly irresponsible. Returning to Karen's question and, in particular, considering her statement that "*obviously she can have the gene in the other breast as well*", provides some further framing

to understand the context in which preventive surgery might be considered. By presenting "*the gene*" as a (possibly location-specific) risk factor, this statement acts to rule out model 6e. That is, the statement serves the pragmatic purpose of restricting the latent causal structure to those possibilities that would be consistent with other models. However, models 6b and 6c allow future cancers to affect the original breast, while model 6d does not. Naturally, the potential benefits of the surgery *S* are dependent on which of these possibilities holds, as illustrated in Figure 6f. From the sisters' perspective, the distinction between 6b and 6c is irrelevant, and so does not enter into Karen's question. However, if either of these is the case, then a bilateral mastectomy is "required", whereas 6d only "needs" a unilateral mastectomy to reduce all risks to (near) zero. This tension is illustrated in 6f, which uses dotted lines to illustrate possible effects of the interventions *T* and *S* (note that the past risk θ_{21} is disconnected from its usual cause, since the intervention *S* clearly cannot influence the past: strictly, the model should differentiate between the value of ϕ_2 pre-*S* and post-*S*, but this would add yet another complexity to an already difficult concept). In summary, Karen's question can be seen to reflect uncertainty about how to understand the causal role of a past treatment of breast cancer. Consideration of the uncertainties regarding various causal relationships implicit in her question is highly relevant to decisions about both her and her sister's ongoing management of increased breast cancer risk.

In light of this discussion, it is not surprising that the counsellor's response is difficult to follow, and appears not to address the content of Karen's question very well. There are a range of variations and adaptations to the complicated causal models we have already considered implicit in the counsellor's talk, which we will not consider here. Focusing only on Karen's question, however, has allowed us to demonstrate some of the ways in which causal representations may be implicit in natural talk, and to consider briefly how such representations may act to frame or constrain subsequent talk and action (e.g., preventive health decisions).

Inferring Causes from Effects

The final extract we present illustrates a more explicit engagement with causation on the part of the speakers. That is, the speakers explicitly attempt to 'model' the cause of a particular occurrence of cancer in an individual. As mentioned previously, explicit engagement with the causal relationship between events and variables is relatively rare and often needs to be inferred from talk. When such explicit engagement does occur, it is clearly of interest in an examination of the causal reasoning apparent in natural language. Extract 5 comes from a session involving a counsellor, a geneticist, and the client. The exchange occurs in the context of a discussion about the difficulties inherent in attempting to definitively assign genetic causes to the occurrence of cancer and the potential benefits of genetic testing.

Extract 5

Geneticist. Tsk. Thee, when we look at cancer overall, maybe 5 to 10% of the people with cancer have an inherited tendency or predisposition. It means that 90 to 95% of people with cancer had bad luck. If you regard something as inevitable as bad luck

Client. mm

Geneticist. and Jane (the counsellor) will go through that (inaudible). Ehm, we recognize that there are some families where there's this inherited tendency to develop certain cancers, and that that is what, we recognise that clinically by looking at family histories. So, there are guidelines if you have three close relatives who have breast or ovarian cancer, the chances are that there is some inherited tendency. The difficulty is that breast and ovarian cancer is sufficiently common that by chance you will come across the occasional family where three women have had those sorts of cancers and there is no genetic tendency. And we have great difficulty in that situation of sorting it out, who has a genetic inherited tendency that we might potentially find out about. Or that we might not. And who has really bad luck but not a bad gene. And the dilemma for us is that we can go down certain paths and do certain genetic studies that might confirm an inherited tendency, but we have no means of excluding an inherited tendency. There's no genetic test for luck.

…

Geneticist. There may be a familial component, {mm} there may be an environmental component, there's certainly a, well, (inaudible) there is a familial component, I just don't know what it is {mm} or how strong it is {mm}. There's a familial component to breaking your leg because you have genes that dictate how strong your bones are. There will be an environmental component, I just have no idea of what it is. And there will be a chance component. And we can't measure that.

Client. So if it was a positive result, a definite positive result

Geneticist. Yes

Client. Would you then talk about testing the children?

Geneticist. Absolutely. Because then we've got something that, you know, we can provide usable information

The nature of the causal models underlying the reasoning in this extract is fundamentally different from that examined in previous extracts. In previous extracts, causal reasoning was concerned with determining what outcome (either cancer or not) will result from various factors (genes, risk, previous cancer, etc.). In contrast, in Extract 5, the causal explanation is concerned with outlining the particular processes (bad luck or bad gene) that may have led to a known outcome (cancer).

---- Figure 7 about here ----

Figure 7a captures the causal framework implied by the geneticist in the earlier part of the extract. It is important to observe here that c (chance or luck) and g (individual genetic factors) are not denoting two factors contributing to a particular (known) outcome but, rather, that they represent two alternative, mutually exclusive causal pathways whereby the outcome may have been achieved. That is, cancer is depicted in this causal model as being a result of either genes or luck, not a combination of the two. Which pathway is the correct one, however, is uncertain and is the subject of the debate.

In contrast, in the latter part of the extract, the geneticist shifts to an additive model in which the outcome is assumed to be the result of a combination of three known factors. Here, in addition to chance and genes, an environmental component (e) is introduced into the causal process. In this model (Figure 7b), rather than the pathway being uncertain, the weighting of each factor in contributing to the occurrence of cancer is unknown (i.e., how much each factor contributes to the eventual outcome of cancer).

In summary, one of the uncertainties commonly dealt with in genetic counselling for cancer relates to the causal mechanisms that have led to particular instances of cancer. Extract 5 illustrates the construction of two similar causal frameworks for this situation: 1) cancer is caused either by a genetic predisposition or by bad luck; 2) cancer has a genetic, an environmental, and a chance component, all of which interact to lead to cancer in a particular individual. These causal maps (Figure 7) are very distinct from those examined previously as the concept of 'risk' has disappeared entirely. In addition, it appears that the stochastic components that were previously bundled into the 'manifestation' part of the chain ($\theta \rightarrow x$) have been extracted and postulated as an alternative causal process ('chance' or 'luck'). It is also worth noting that although the models represented in 7a and 7b are conceptually quite distinct, the speaker in Extract 5 was able to move between them without much difficulty (i.e., no conversational tension or 'trouble' is evident in the transcripts). Thus, the two implied models should not necessarily be seen as contradictory, but as complementary to the larger purpose of communicating an adequate level of understanding of familial cancer risk to clients.

Discussion

This paper has been concerned with illustrating some of the complex ways in which causal reasoning is embedded in natural language. Causal language provides people with a powerful tool with which to engage with questions about the relations between events, people and ideas. In a recent review, Sloman (2005, p21) notes that "In everyday language, causes and effects assume various roles. We say drugs cause addiction, sparks cause fire, and guns cause death". In order to attach meaning to this kind of statement, he suggests (p24) that

> To say that A caused B seems to mean something like the following: A and B both occurred, but if event A had not occurred (and B had no other sufficient

> causes), B would not have occurred either. The critical point is that a causal relation doesn't merely imply that events happened together but that there's some generating mechanism that produces an event of one time when engaged by an event of another type. So in some other world in which the mechanism had not been engaged by the cause, the effect would not have resulted. This is what distinguishes a causal relation from a mere correlation.

There is a great deal of evidence suggesting that people learn such causal relations in the environment through both 'pure' observation and planned interventions (e.g., Steyvers et al, 2003). Indeed, most formal models of causal reasoning developed so far, such as Cheng's causal power model (Cheng, 1997), are essentially restricted to cases of direct observation and intervention. However, broader discussions about causal reasoning (e.g., Sloman, 2005) suggest that we undergo similar learning processes in purely linguistic interactions. It thus seems desirable to complement existing knowledge of causal reasoning derived from experimental studies with analyses of how causal structures are expressed in naturalistic language (i.e., in settings not contrived by the researcher). To this end, this study has served to present an analysis of real-world causal talk. It is also interesting to note that while our methodological approach to causal reasoning differs significantly from typical quantitative experimental studies, our analysis shows that there is much similarity in the central ideas utilised in more traditional studies investigating causal reasoning (e.g., constraints of domain theory, reasoning based on interventions, etc.). One of our hopes is therefore that, beyond being interesting in their own right, studies of the kind we presented here can help to motivate further developments in experiments and theories about causal reasoning.

Although we have not yet explicitly investigated the way in which individuals take up causal relations latent in language, we argue that such causal models constrain (or, at very least, frame) subsequent ways of understanding causal relationships, as well as decisions and actions. That is, one of our central claims is that people learn from the causal representations of the world that are implicit in talk. We suggest that this is particularly so for those causal relationships between events, objects and outcomes about which it is impossible for most people to learn explicitly through observation (e.g., the link between genes and cancer). Our approach to studying causal reasoning, therefore, contributes to the literature on causal reasoning by demonstrating that: (1) complex causal relationships can be found in naturally occurring talk; (2) the representations implicit in talk are often varied and even contradictory; and (3) qualitative data constitute a valuable source of information about how individuals use causal models in naturalistic settings and it is possible to analyse such data with a high degree of rigour. One important shortcoming of the present study is that we do not yet know how and to what degree causal models embedded in language are taken up by listeners, or how their causal beliefs are modified as a result. One way in which this could be studied in the context of the domain of familial cancer is to apply our approach, which we illustrated through analysis of counsellors' talk, to clients' talk. Thus, to study the uptake of causal reasoning by clients, it is simply a matter of conducting a systematic

mapping of causal models implicit in clients' talk and comparing these to the causal structures implicit in the health professionals' talk. As it was our purpose in this paper to illustrate the principles of our analysis, space constraints made a more systematic study of this kind infeasible. However, we see the uptake of causal structures embedded in naturally occurring talk as an important area of future research.

A number of additional observations can be made of this study and our approach to causal reasoning. First, we should point out that the extracts presented in this paper do not represent a comprehensive treatment of the kinds of causal structures implied or conveyed to clients in this setting. Rather, they represent a 'sample of convenience' that we selected to illustrate a range of causal maps that can be inferred from spoken language. The reason for this approach is that our aim was not to document the process of genetic counselling, but to illustrate a novel approach to the study of causal reasoning. The domain of familial cancer is too complex for us to be able to provide an exhaustive catalogue of the causal structures employed in this setting in just one paper.

Second, we take a descriptive approach to the study of causal reasoning processes. We agree with Ahn and Kalish (2000) who maintain that normative models cannot account for everyday causal reasoning processes as most naturalistic domains are simply too complex. In addition, our approach to causal reasoning also emphasises the degree of ambiguity inherent in the stimuli that underlie the formation of causal beliefs. In particular, there is a degree of subjectivity in identifying causal structures that are implicit in language (particularly spoken language). This element of subjectivity is reflected in our own analysis. Not everyone might agree, for instance, that the speaker in Extract 3 favours a symptomatic over a causal model connecting genes and cancer. However, we see this subjectivity as supporting, rather than detracting from our argument. If even careful analysis of these conversations cannot determine one definitive model that is being implied, then how can clients, who have to make up their minds during the course of the conversation, be expected to infer the 'correct' model?

Third, although our focus was less on genetic counselling than on promoting a new approach to causal reasoning, there are practical implications to this kind of analysis for clinical practice. Causal reasoning has been shown to be an important factor in responses to risks. In a study by Senior, Marteau and Peters (1999) on familial hypercholesterolaemia (FH), it was shown that responses to screening varied depending on whether a positive test result was perceived as detecting raised cholesterol or detecting a genetic problem. In particular, parents of children who tested positive for FH were found to cope much better when the cause of the problem was seen to be high cholesterol, rather than genetic. Since these perceptions are formed through communication (and not 'direct' observation), studies of causal reasoning in naturalistic settings may also have significant practical benefits to offer.

Finally, we do not present our case as a way of arguing against conducting laboratory experiments on causal reasoning. Rather, our point is simply that everyday causal reasoning is far more complex than many experimental studies imply. Many people hold causal beliefs about domains in which they have had no direct experiences of causes and effects. These causal beliefs are formed as the result of some form of

communication. It stands to reason, therefore, that to understand causal reasoning, these processes of communication need to be studied. While we maintain that our approach adds another dimension to existing studies on the topic, we have not yet attempted to generate a more comprehensive theoretical account of causal reasoning. We are very far from being able to do so. However, we do believe that we have presented certain complexities that cannot be ignored and we present our arguments as a starting point to further discussion.

References

Ahn, W., & Kalish, C. W. (2000). The role of mechanism beliefs in causal reasoning. In F. Keil & R. A. Wilson (Eds.), *Explanation and cognition*.(pp. 199-226). Cambridge, MA: MIT Press.

Cheng, P. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53*, 134-140.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* Oxford, UK: Oxford University Press.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111* (1), 1-31.

Grice, P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (eds). *Studies in Syntax,* vol 3. (pp. 41-58). New York: Academic Press.

Griffiths, T. L., Baraff, E. R., & Tenenbaum, J. B. (2004) Using physical theories to infer hidden causal structure. *Proceedings of the 26th Annual Conference of the Cognitive Science Society.*

Griffiths, T. L., & Tenenbaum, J. B. (2005) Structure and strength in causal induction. *Cognitive Psychology, 51,* 354-384.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science.* Cambridge, UK: Cambridge University Press.

Jefferson, G. (1990). List construction as a task and resource. In G. Psathas (Ed.) *Interaction Competence* (p.63-92). Lanham: University Press of America.

Lagnado, D. A., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology, 30*(4), 856–876.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2006). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal Learning: Psychology, Philosophy, and Computation*. Oxford, England: Oxford University Press.

Lynch , E., & Medin, D. (2006). Explanatory models of illness: A study of within-culture variation. *Cognitive Psychology, 53*, 285-309.

Navarro, D. J., Griffiths, T. L., Steyvers, M. & Lee, M. D. (2006). Modeling individual differences using Dirichlet processes. *Journal of Mathematical Psychology, 50*, 101-122.

Neisser, U. (1981). John Dean's memory: A case study. *Cognition , 9*, 1-22.

O'Doherty, K. C. (2006). Risk communication in genetic counselling: A discursive approach to probability. *Theory & Psychology, 16* (2), 225-256.

O'Doherty, K. (2007). Implications of conflicting definitions of probability to health risk communication: A case study of familial cancer and genetic counselling. *Australian Health Review, 31*(1), 24-33.

O'Doherty, K. C. & Suthers, G. K. (2007). Risky communication: Pitfalls in genetic counselling about risk, and how to avoid them. *Journal of Genetic Counseling*. …

Over, D. E., Hadjichristidis, C., Evans, J. S. B. T., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54*(1), 62-97.

Pearl, J. (2000). *Causality: Models, Reasoning & Inference.* Cambridge, UK: Cambridge University Press.

Rehder, B. (2003). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology, 29* (6), 1141-1159.

Rouder, J. N., Sun, D., Speckman, P. L., Lu, J., & Zhou, D. (2003). A hierarchical Bayesian statistical framework for skewed variables with an application to response time distributions. *Psychometrika, 68*, 587-604.

Senior, V., Marteau, T. M., & Peters, T. J. (1999). Will genetic testing for predisposition for disease result in fatalism? A qualitative study of parent's responses to neonatal screening for familial hypercholesterolaemia. *Social Science & Medicine, 48*(12), 1857-1860.

Sloman. S. A. (2005). *Causal Models: How People Think about the World and its Alternatives.* Oxford University Press.

Sloman, S. A., Love, B. C., & Ahn, W.-k. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22*(2), 189-228.

Sloman, S. A. & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29*, 5-39.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J. & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27,* 453-489.

Von Mises, R. (1957). *Probability, Statistics & Truth.* London: Allen & Unwin.

Wetherell, M., Taylor, S., & Yates, S. J. (Eds.). (2001). *Discourse as Data: A Guide for Analysis*. London: Sage.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition, 88*, 1-48.

Wolff, P., Song, G., & Driscoll, D. (2002). *Models of Causation and Causal Verbs.* Paper presented at the 37th Meeting of the Chicago Linguistics Society, Chicago.

Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics, 5*, 161-215.

Figures

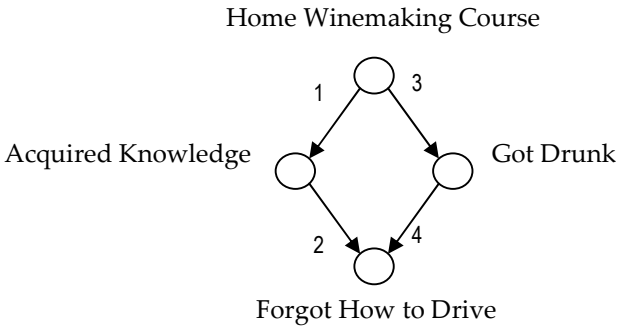


Figure 1: A directed acyclic graph that represents two competing explanations for Homer Simpson's poor driving.

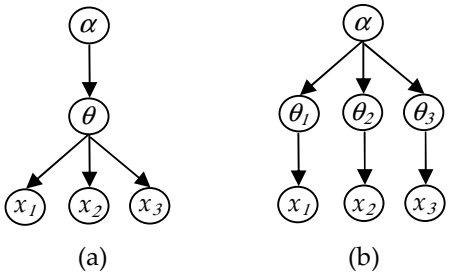


Figure 2: A shared risk model (a) versus an *i.i.d.* risk model (b). In both graphs, α denotes some shared background (implicitly corresponding to a "normal" genetic condition and environment), θ denotes a risk of developing cancer, and x denotes the eventual (binary) outcome. In both panels, the different subscripts indicate different people or 'instances'.

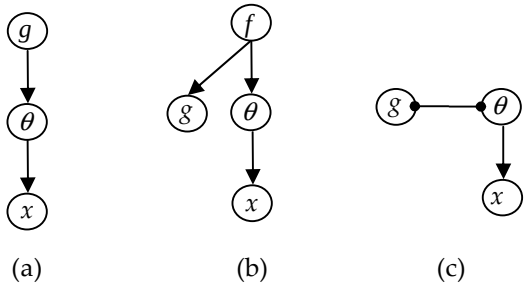
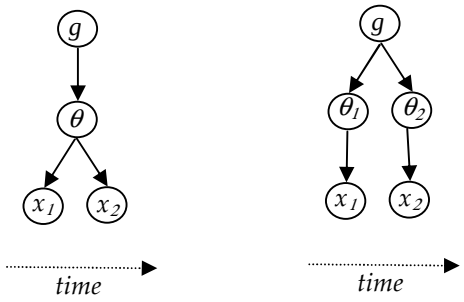
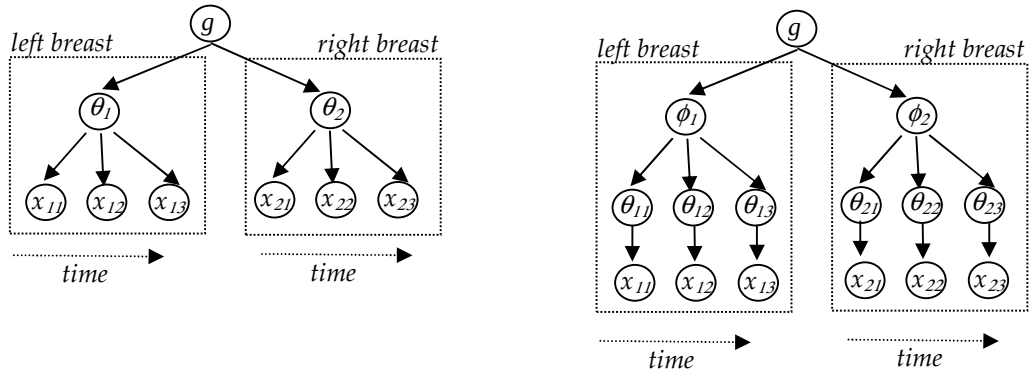


Figure 3: Elaborating the causally relevant background, using three simple models for genetic covariance. In all models g denotes the state of the "genetic condition", f denotes unknown (or unspecified) causal factors, and as before θ and x denote risks and outcomes. In a *strictly causal model* (a), the genetic condition is explicitly viewed as the cause of the risk. In a *strictly symptomatic model* (b), the genes and risks share a common cause. Finally, in a *correlational model* (c), the causal connection between genes and risks is left deliberately ambiguous.



(a) (b)

Figure 4: Two potential causal models for the appearance of cancer at different points in time, though expressed at the same location. Here, g represents underlying genetic factors, and θ and x denote risks and outcomes. Subscripts indicate specific outcomes (and, in panel (b), risks) at discrete points in time.



(a) (b)

Figure 5: A richer representation of multiple risks and multiple cancers within a single individual. In the subscripts ij, i denotes location (left or right breast), j denotes outcomes (and risks for panel (b)) at different points in time. The left panel thus incorporates differential risks for multiple locations, but not multiple time points. That is, while the model allows for different outcomes over time, risk remains constant over multiple time points. In contrast, the right panel allows the risk to change over time in each breast, according to some (unspecified) breast-specific factor ϕ.

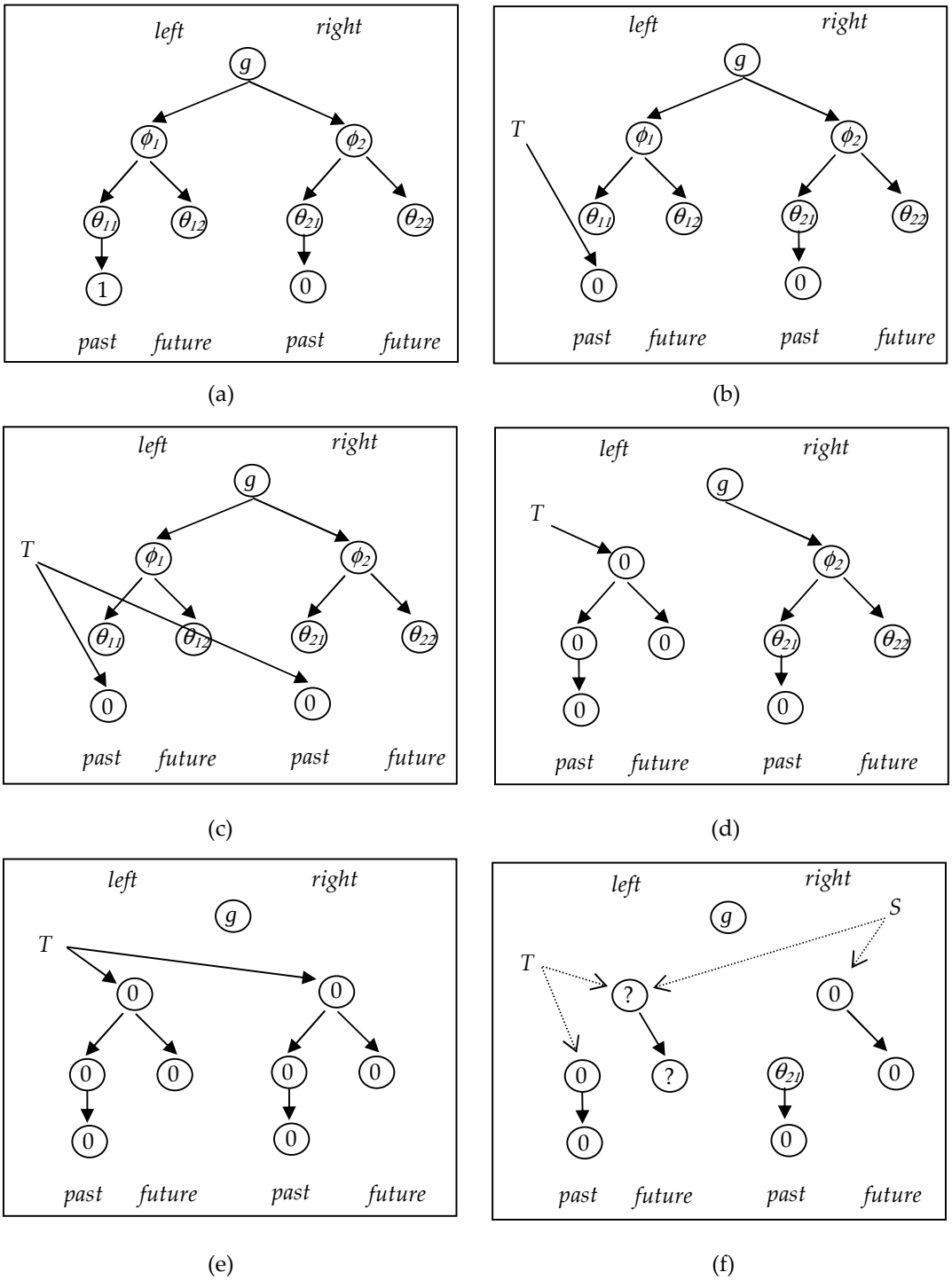


Figure 6: Developing a model to cover Karen's question.

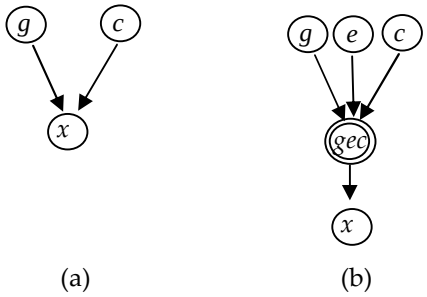


Figure 7: Gene or chance? Or an additive model? Panel (a) illustrates the representation of *g* (genetic factors; "bad gene") and *c* (chance; "bad luck") as competing causal pathways for the occurrence of cancer. Panel (b) illustrates an additive model in which *g*, *c*, and *e* (environmental factors) all contribute to the outcome.