# Causality in Decision Making

York Hagmayer, University of Goettingen, Germany

Philip M. Fernbach, University of Colorado at Boulder, USA

Corresponding Author

York Hagmayer

Institute of Psychology

University of Goettingen

Gosslerstr. 14

37073 Goettingen

Germany

Phone: +49 551 398293

Email: yhagmay@gwdg.de

Abstract

Although causality is rarely discussed in handbooks on decision making, decisions often depend on causal knowledge and causal reasoning. Our goal is to review what is known about how people integrate causal considerations into their choice processes. We first introduce causal decision theory, a normative theory of choice based on the idea that rational decision-making requires considering the causal structure underlying a decision problem. We then provide an overview of empirical studies that explore how causal assumptions influence choice and test predictions derived from causal decision theory. Next we review three descriptive theories that each integrates causal thinking into decision-making in a different way: The causal model theory of choice, the story model of decision-making, and attribution theory. We discuss commonalities and differences between the theories and the role of causality in other decision making theories. Finally, we conclude by posing some open questions and challenges that lie ahead for research on the role of causal reasoning in decision making.

Keywords: Causal decision making, causal model theory of choice, causal narratives, story-model of decision making, attribution theory

## 1. Introduction

Decisions generally concern potential actions (e.g., going on a diet) or objects that afford actions (e.g., choosing groceries to prepare for dinner). There are many ways these decisions are made (see Koehler & Harvey, 2007, for an overview). Some decisions are guided by social, moral or legal norms (e.g., not to lie), some are guided by emotions (e.g., avoidance based on disgust), but many are based on the outcomes that we expect to result from actions afforded by the decision options. For example, we choose to diet because we expect to lose weight, and we decide on a set of ingredients at the grocery story because we believe they will combine to make a good meal. Which outcomes actually result, depends on underlying causal mechanisms. Whether dieting results in weight loss, depends on a complex network of physiological and psychological processes. For instance, dieting may cause cravings for sweet and fatty food, resulting in binge eating, disturbed eating patterns, and sometimes even in weight gain (Howard & Porzelius, 1999). Causal knowledge about the system in question can help us to make more adaptive decisions. Knowing that dieting sometimes triggers binging allows a dieter to structure his or her behavior and environment to avoid this negative consequence.

In order to predict the consequences of different options, it is sometimes also necessary to analyze the causes of the current situation. Take a general practitioner, who is consulted by an emaciated patient. Serious underweight can be caused by many factors, including anorexia, diabetes, and cancer. Depending on the cause, the very same intervention can have different consequences. A high calorie diet could be beneficial for a cancer patient but harmful for an untreated diabetic. Sometimes diagnosing the causes of a situation is sufficient to predict which decision option will yield the best consequences. Consider a physician who has to decide on how to manage a patient's cold. By determining whether the symptoms are due to a viral or bacterial infection, the decision becomes obvious.

These examples show that causal knowledge supports decision making in two ways: By allowing us to predict the consequences of different actions under the given circumstances and by helping to make diagnoses that suggest which interventions will be effective. Thus causal knowledge and inferences based on this knowledge seem to be critical for good decision-making.

The importance of causality to advantageous decision-making is potentially problematic because research into causal explanations has shown that people often have only rough, skeletal knowledge about causal mechanisms (Keil, 2008). People often do not know how everyday objects like bicycles or can openers work (Lawson, 2006; Rozenblit & Keil, 2002), or how favored public policies will lead to beneficial consequences (Fernbach, Roger, Sloman & Fox, 2013). Although some people have rather elaborate lay theories about certain domains like medicine and biology, these theories are often inconsistent with scientific consensus and tend to be incomplete (Furnham, 1988). Therefore people's causal knowledge only allows for very rough, and sometimes incorrect predictions of consequences. Given that our causal knowledge is incomplete or sometimes wrong, it might be futile or even harmful to try to base decisions on causal considerations.

The remainder of this chapter is organized into four parts: First, we explore the normative question of whether the causal structure underlying a decision problem should be considered when making a decision, and we describe Causal decision theory, which claims that causal considerations are necessary for rational decision making. Second, we present evidence from empirical studies investigating how people use causality in decision-making. Third, we introduce three descriptive models from cognitive psychology, which assume that causal reasoning is central to decision making. Finally, we end by discussing open questions and challenges that lie ahead for causal decision making research.

## 2. Should decisions be based on causal considerations?

Economists and psychologists have argued that rational decision makers' choices must conform to the recommendations of *Expected Utility Theory* (Savage, 1954; von Neumann & Morgenstern, 1944). The theory distinguishes between actions ($A_i$), outcomes ($O_{ji}$), utilities of outcomes ($U[O_{ji}]$), and the "state of the world" ($S_j$), which encompasses all the variables that affect the outcome beside the action. Figure 1 illustrates the basic structure of the decision situation and shows a table demonstrating an example. The theory assumes that every outcome can be assigned a utility, which represents the usefulness of the outcome for the decision maker. For simplicity assume that the costs of the action (i.e., the time and resources required to perform the action) are already included in the utility of the outcome. The theory also assumes that every outcome can be assigned a probability $P(O_{ji})$, which represents the uncertainty that a particular outcome will result after the actions has been taken. Outcomes can be uncertain for several reasons: (a) the state of the world, which determines the outcome, is not known for sure $P(S_j) < 1$, and/or (b) it is not known for sure whether the state of the world in conjunction with the action taken, is followed by the outcome $P(O_{ji}|S_j, A_i) < 1$. Both probabilities reflect the current lack of knowledge or evidence concerning the presence of the outcome given a particular action. To make a decision, the expected utility (EU) of every action $A_i$ needs to be calculated by multiplying the probability of each possible outcome with the utility of the respective outcome and summing up the products. Formally:

$$EU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$$

$$= \sum_{ji} P(O_{ji}|S_j, A_i)*P(S_j) * U(O_{ji})$$

The theory dictates that the action with the highest expected utility should be performed. That is, choice should be based on the *principle of maximizing expected utility* (MEU-principle).

The table in Figure 1 provides an example of how expected utilities are calculated. Imagine that a forty-year-old woman has to decide whether to start breast cancer screening (Action 1) or not (Action 2). The state of the world (S) is either the presence of breast cancer ($S_1$) or the absence of cancer ($S_2$). The major outcomes are to die early from cancer ($O_1$) or to live long ($O_2$). Obviously the utility of a long life is much higher than the utility of dying early. How big the difference is, depends on the decisions maker's happiness in life. The screening procedure has some cost, because it takes time, creates anxiety, etc. Therefore, the utility of the outcomes after screening are lower than the corresponding outcomes without screening. The outcome with the lowest utility is usually assigned zero, all other outcomes are assigned utilities that reflect how beneficial they are for the decision maker. In the example shown, we assume that a long life has a much higher utility than the costs (e.g., time, pain, anxiety) of cancer screening.

The benefit of screening results from the early detection of cancer. When cancer is detected in an early stage, it can be treated more successfully. Hence the probability of living long is higher for women who are screened. Research indicates that the lifetime probability of a woman getting breast cancer is roughly 12% ($P[S_1]=.12$; www.cancer.gov), roughly 1 in 4 women that contract breast cancer die from it, which means that overall about 3% of women die from breast cancer. The risk to die from cancer is reduced through screening starting at age 40 by about 20%, that is, from 3% to about 2.4%. Based on these probabilities and the utilities chosen to illustrate the example in the table, screening has a slightly higher expected utility than no screening. Therefore, screening should be chosen according to the principle of maximizing expected utility. Note that the decision could change if the assigned utilities were changed.

>> Insert Figure 1 about here <<

Expected utility theory has intuitive appeal and is considered the gold standard normative decision-making theory. Proponents of *Causal decision theory*, however, criticize expected utility theory, because it does not explicitly consider the causal basis of outcome probabilities (Lewis, 2006; Maher, 1987; Nozick, 1993; Skyrms, 1982; Joyce, 1999). Proponents of causal decision theory point out that a probabilistic relation between an action and an outcome sometimes reflects a spurious, non-causal relationship. According to causal decision theorists, this distinction is crucial for good decision making because an action only increases the decision maker's utility if the action causally affects the outcome, but not when it is spuriously related. Thus, decisions should not be based on merely on observed statistical relationships, but based on the underlying causal structure. Nozick (1993) elaborated this idea by proposing that decisions should be based on *causally expected utilities* (CEU) rather than *evidentially expected utilities* (EEU).

To make this more concrete consider the following example, which is related to the breast cancer example above. Women who worry a lot about breast cancer tend to have a greater likelihood of dying from breast cancer than women who do not. There is little evidence that worry causally relates to death. Instead this correlation is spurious. Those who worry more, are more likely to actually have cancer and hence have a higher probability of dying. (For instance, worry may be triggered by unusual signals from the body). As a consequence, it does not make sense to stop worrying in order to increase the chances of survival.

Actions, the state of the world, and the resulting outcome may be related through various causal structures. Figure 2 shows four possible causal structures. First, actions may generate the outcome (Figure 2a). For example, dieting causes weight loss. Second, actions may not cause an outcome, but they may enable other variables, which are part of the state of the world, to generate the outcome (Figure 2b). For example, buying a lottery ticket does not cause us to win, but it enables us to win. Our purchase has no influence on which ticket is drawn as the winner. It is the draw, which determines whether we win a lot of money. Third, actions may not directly affect the outcome, but instead indirectly affect it by influencing the state of the world (Figure 2c). For example, buying health insurance does not make us healthy. But having health coverage (the resulting state of the world) causes us to get better access to health care, which in turn results in better health. Finally, as in the example above, actions and outcomes may be spuriously related (Figure 2d). This is the case when the action and the outcome are caused by some other factor. Due to this common cause the action and the outcome are statistically but not causally related to each other.[1]

Figure 2 also outlines formalizations for calculating causal expected utilities (CEU) given the different causal structures. As the formalizations show, the causal expected utilities are calculated differently depending on the underlying causal structure. In the first three cases causal expected utility can be calculated by $CEU(A_i) = \sum_{ji} P(O_{ji}|A_i) * U(O_{ji})$, hence by taking into account the statistical dependence between the action and all possible outcomes. However, it would be wrong to do so when the action and the outcome are spuriously related. In this case, the outcome and the action are independent of each other if the action is deliberately chosen, formally $P(O_{ji}|A_i) = P(O_{ji})$.

[1] Figuring out whether an action and an outcome are causally or spuriously related may not be easy. The famous statistician Sir Ronald Fisher argued for many years that the reliable statistical relation between smoking and lung cancer might be due to genetic factors (e.g., Fisher, 1958). Recently researchers showed that some of the reliable statistical relations between music practice and musical abilities are probably due to genetic factors (Mosing, Madison, Pedersen, Kuja-Kalkola, & Ullen, 2014).

Therefore the causal expected utility has to be calculated as $CEU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$. This equation takes into account that taking the action will not change the probability of the outcome.

>> Insert Figure 2 about here <<

In order to predict the likelihood of the outcome correctly, it is important to distinguish between choosing the action and observing the action. When there is a spurious relation between an action and an outcome, observing the action is predictive of the outcome, but choosing the action does not increase the likelihood of the outcome. Consider again the example of worry and breast cancer. If you observe someone worrying about breast cancer, this should raise your subjective probability that she has breast cancer, and, as a consequence, your subjective probability that she will die from breast cancer. But this is not what is relevant for decision making. For decision making the consequences of making a choice are relevant. If she were to decide to stop worrying, her chances of dying from breast cancer would not be affected. Figure 3 illustrates this point. The left panel shows what can be inferred about the likelihood of dying from cancer from observing worry. The right panels shows what will happen if the same person decides to worry or not to worry. Deciding entails that the action is now determined by choice. It is no longer determined by the presence or absence of breast cancer. By virtue of choosing the action, the action becomes independent of breast cancer. Worry is no longer predictive of breast cancer, and therefore worry is no longer predictive of the likelihood of dying from breast cancer.

The distinction between observing and choosing an action can be formally captured by a distinction made in *causal Bayes net theories* between observing the value of a variable and an intervention that sets the variable to a specific value. (Pearl, 2000; Spirtes, Glymour, & Scheines, 1993/2000; Rottmann, this volume). Interventions render the intervened-on variable independent of all other variables (except the variable's effects (Pearl, 2000). Thus the intervened-on variable becomes independent of its usual causes. Observed variables, by contrast, are probabilistically related to their causes. Deliberate choices are like interventions in that they exogenously determine the action and render it independent of other factors ((Hagmayer & Sloman, 2009; Sloman, Fernbach, and Hagmayer, 2010). For instance, deliberately choosing not to worry renders worry independent of its normal causes (e.g. the presence of cancer).

>> Insert Figure 3 about here <<

To summarize, causal decision theory argues that knowledge of the statistical relationships between an action and a desired outcome is sometimes insufficient for making the best choice. In order to make an optimal decision, decision-makers have to understand the causal structure that relates actions, potential states of the world andoutcomes. Moreover, choice should be treated as an intervention, not an observation. Based on assumptions about the underlying causal structure causal expected utilities can be calculated which allow the decision maker to identify and choose the action that increases the probability of obtaining the desired outcomes the most. These points are not necessarily inconsistent with expected utility theory. Instead they show that reasonable utilities can only be generated if one has knowledge of the causal structure relating the chosen action and the desired outcome. Thus, causal knowledge is central to rational decision-making.

## 3. Do people engage in causal decision making?

Causal decision theory is a normative theory that specifies how a rational decision maker should determine the utility of decision options. From a psychological perspective we can ask whether it describes how people actually make decisions. If Causal decision theory is interpreted as a descriptive theory it makes two basic predictions. First, decision makers should analyze the causal structure underlying a decision problem in order to be able to predict the consequences resulting from the available options. Second, decision makers should base their decisions on the causal expected utilities of the given options. This implies that, when predicting consequences, they should treat their choice as an intervention, which entails that actions become independent of other causal factors that also affect the outcome. We will explore whether these predictions are supported by empirical evidence in the following sections.

### 3.1 Do people analyze the causal structure underlying a decision problem?

As we have explained, people should analyze the underlying causal structure to understand the consequences of different decision options. For example, when choosing between different treatments for depression, it is important to know whether the depression is a reaction to some other medical or mental problem. If it is a reaction, addressing the other problem will be more effective for ameliorating depression than addressing depression directly. Therefore, the causes of depression have to be analyzed before the consequences of the different treatment options can be predicted.

There is good evidence that people spontaneously search for causes when they are confronted with an unexpected, threatening or norm-violating event (Weiner, 1985; Kahneman & Miller, 1986; Kahneman, 2011). In these cases they use the available information and background knowledge to infer the causes of what happened. Research on naturalistic decision making reports similar findings. When decision makers recognize a situation as a familiar decision problem, they know what to do. But when the situation is novel, they initiate a causal analysis (Klein, 1998; Zsambok & Klein, 1997).

Other research has explored whether decision makers analyze the causal structure connecting actions and outcome. Hagmayer and Sloman (2009, Experiment 2) presented participants with a statistical relation between an action and an outcome (e.g., "people, who watch movies in their original language, speak better English than people, who do not"), and asked whether they would recommend the action to a friend who is interested in achieving the outcome. After participants made their recommendation, they were asked whether the relation was due to a direct causal relation between the action and the outcome or a common cause influencing both the action and the outcome. Participants who believed in a common cause did not recommend taking the action, while people who believed in a direct causal relation did. The same finding resulted for relations, which participants considered plausible and implausible in a pretest. This indicates that participants probably considered the underlying causal structure before making their choice and did not rely merely on the plausibility of the given relation.

Hagmayer and Meder (2013) investigated whether participants analyze causal structure when making repeated decisions. Participants were asked to maximize the vaccine produced by batches of genetically modified bacteria. To do so, they could choose between different trigger substances. In the "common cause" condition, the trigger with the highest payoff affected two relevant genes directly. In the other "chain" condition, the same trigger affected only one gene, which in turn activated the second gene. Participants not only learned which trigger maximized the outcome, many also learned about the causal structure connecting actions to outcomes. Importantly, assumptions about causal structure strongly affected later decisions when new options became available. There is, however, conflicting evidence from other studies on repeated decision making and control of dynamic systems, which show many participants do not learn about causal structure (Hagmayer et al., 2011; Osman, 2010; see also Osman, this volume).

In sum, the first prediction of causal decision theory seems to be supported. People analyze the causal structure underlying a decision problem when the situation is sufficiently novel to warrant a causal analysis and when they have sufficient background knowledge to infer the underlying causal structure.

## 3.2 Do people base their decisions on causal or evidential expected utilities?

According to causal decision theory people should base their decisions on causal expected utilities rather than evidential expected utilities. To do so, they have to consider the causal relation between an action and an outcome and not merely the statistical relation, which could be spurious. Several experiments have investigated this prediction (Hagmayer & Sloman, 2009; Robinson, Sloman, Hagmayer, & Hertzog, 2010). Robinson and colleagues (2010) presented participants with a number of economic games based on Prisoner's Dilemma. In one scenario participants were asked to imagine being a currency trader at the bank of Japan having to decide whether to buy dollars or euros. A competitor would do the same. Depending on their choices, different profits would result. The expected payoffs of the two choices were presented as a table (see Table 1). They were also told that in the past they and their competitor made the same decision 90% of the time. Based on this information, the evidential expected value of buying dollars would be substantially higher ($905 million) than the expected value of buying euros ($210 million). Disregarding the information about statistical relations in the past, however, it would be better to buy euros regardless of the competitor's decision (1.2 billion vs. 1 billion when the competitor buys dollars, 100 million vs. 50 million when the other person purchases euros). Participants were also told about the underlying causal structure. In one condition they were told that the relation between their own and their competitor's decision was due to a common cause: Both base their decision on the same economic data, which has caused them to make the same decision in the past. In a second condition, they were told that the relation was due to the competitor waiting to see the participant's decision before deciding. Hence, the participant's decision directly causes the competitor's choice. In a third control condition, no information about the underling causal structure was provided. Respective assumptions were assessed after participants made their choice. The results showed an effect of causal beliefs upon decisions. Participants were more likely to buy dollars when they believed their choice was a direct cause of their competitor's decision than when there was a common cause. These findings and the results of other experiments (DeKwaadsteniet et al., 2010; Flores, Cobos, Lopez, & Godoy, 2014; Yopchick & Kim, 2009; Hagmayer & Sloman, 2009) show that decision makers tend to maximize causal expected utility rather than evidential expected utility.

>> Insert Table 1 about here <<

Causal decision theory prescribes that decision makers should differentiate between observing an action and choosing an action when predicting the consequences of their choice. They should consider their choice to be an intervention, which renders the chosen action independent of other causes of the desired outcome. The findings described above were consistent with this prediction, but did not directly investigate it. Other studies have investigated more directly how decision makers conceive of their choice.

Studies on agency have shown that deliberate decision makers perceive themselves as free agents responsible for their actions (DeCharms, 1968; Langer; 1975; Botti & McGill, 2011). Decision makers often deny being forced or unconsciously influenced by other factors, even if they objectively are affected by them (Ariely & Norton, 2008; Wegner, 2002; Bargh & Chartrand, 1999). Hence, it seems that decision makers conceive of their choice as independent of other factors.

Hagmayer and Sloman (2009) directly investigated whether people equate choices with interventions. In one experiment (Hagmayer & Sloman, 2009, Experiment 3) participants were asked to predict the likelihood of an outcome given than an action is chosen, enforced by another person or machine, or observed. Assumptions about the causal structure underlying the statistical relation between the action and the outcome were also manipulated. The relation was either introduced as being due to a direct causal relation or as being due to a common cause affecting both action and outcome.  For example, participants were told that research has shown that of 100 men who help with the chores 82 are in good health, while only 32 of 100 not helping with the chores are. It was either explained that doing the chores provides additional exercise every day (direct cause condition) or that men who are concerned about equality issues are also concerned about health and therefore are likely to help with the chores and eat healthier food (common cause condition). Then they were asked to estimate the likelihood that a person will be in good health, if (i) she decides to do the chores, (ii) is forced by somebody else to do it, or (iii) is observed to do the chores. Participants made the same predictions in the first two conditions (deliberate choice and external force). In these conditions participants expected to find good health if they assumed a direct causal relation between doing the chores and health, but not when they assumed a common cause. By contrast, in condition 3 when the action was observed, participants expected good health regardless of causal structure. In the common cause condition, this finding indicates that participants inferred the presence of the common cause from the observed action and in turn expected the outcome. Experiment 4 of Hagmayer and Sloman (2009) showed that participants in fact derive different inferences about a common cause from chosen and observed actions.

There is, however, some conflicting evidence. There are studies which show that people sometimes violate the logic of causal decision theory by choosing options that reduce causal expected utilities. Self-deception is probably the most prominent example (Quattrone & Tversky, 1984; Mijovic-Prelec & Prelec, 2003; Fernbach, Sloman, & Hagmayer, 2014, Sloman, Fernbach & Hagmayer, 2011). For example, in Quattrone and Tversky's (1984) study, participants were told that pain tolerance indicates whether someone has a strong or weak heart. In one condition high tolerance was purportedly indicative of a good heart, while in the other condition the opposite was true. Participants told that high tolerance indicated a strong heart tolerated a painful task (holding one's hand in cold water) longer than those told the opposite. Since pain tolerance is a consequence, not a cause of heart type, participants who increased their tolerance increased their pain without generating positive causal consequences. Thus they violated the principle of maximizing causal expected utility. Presumably they did it to signal to themselves they have a strong heart. However, since the behavior violates causal logic it provides no true evidence about heart type. This is why cases like these indicate self-deception. Related findings come from research on self-handicapping (Urdan & Midgley, 2011).

Taken together, the empirical evidence supports the predictions derived from causal decision theory, although there are some cases where decision makers violate these predictions.

Box 1: Does causal knowledge affect decisions?

There are many ways in which causal beliefs may affect decisions. First, causal beliefs may alter expectancies of outcomes, that is, the subjective probability that an outcome will result from a certain action. For example, research on gambling has shown that many pathological gamblers have an "illusion of control" (Langer, 1975; Langer & Roth, 1975), a false belief that their actions have an influence on their chances of winning (Sylvain, Ladouceur, & Boisvert, 1997). Interestingly, there is good evidence that changing these faulty causal beliefs is an effective treatment for pathological gambling (Ladouceur, Ferland, & Fournier, 2003). In the health domain it has been shown that beliefs about control are the best predictors of patients' choices (see Baines & Wittkowski, 2013; Lobban, Barrowclough, & Jones, 2003, for reviews). People who believe that professional treatments are able to control their illness are more likely to seek respective help and adhere to treatments. People who believe in personal control engage more in active coping and better health behavior (Lobban et al., 2003; Baines & Wittkowski, 2013). Different causal beliefs also tend to result in different health outcomes. People with strong beliefs in personal and treatment control tend to have a better prognosis and experience less distress than people, who believe that their illness cannot be controlled (Lobban et al., 2003; Baines & Wittkowski, 2013; Stockford, Turner, & Cooper, 2007).

Second, causal beliefs may affect which information people search for before making a choice. For example, in research using the *active information search paradigm* decision makers were asked to choose between different options in order to achieve a limited number of goals (Huber, Huberm & Bär, 2011; Huber, Wider & Huber, 1997). In this paradigm decision makers have to actively search for the information they consider relevant by asking the researcher, who provides an answer in a standardized way. They are free to collect as much information as they like before making a final decision. The results show that decision makers looked for the causal consequences resulting from the available options, first positive consequences and then negative ones. They were also interested in exploring additional actions that may counteract possible negative consequences arising from the different options ("*risk diffusing operators,"* Huber et al, 1997).

Third, causal beliefs may affect how decision makers weigh different pieces of information when predicting outcomes before making a choice. Research on cue-based decision making has shown that participants prefer to look up cues that are known causes of an outcome and they weigh these cues more than other equally predictive, but not causally connected cues (Garcia-Retamero & Hoffrage, 2006; Garcia-Retamero, Wallin, & Dieckmann, 2007, Müller, Garcia-Retamero, Cokely, & Maldonado, 2011; Müller, Garcia-Retamero, Galesic, & Maldonado, 2013). These findings are consistent with research showing that causal beliefs can bias subjective probabilities (Tversky & Kahneman, 1980; Koslowski, 1996; Chapman & Chapman, 1969, but see Krynski & Tenenbaum, 2008, for evidence that causal beliefs may also de-bias probability estimates).

Fourth, causal beliefs may affect which variables are targeted through an intervention when the outcome cannot be directly manipulated. For example, we cannot directly manipulate our body weight, but we can change its causes (e.g., through physical activity or diet). When people believe that many variables have a direct or indirect impact on the desired outcome, they tend to judge interventions that address strong causes as more effective than interventions targeting other, weaker causes. For example, Ahn, Proctor, and Flanagan (2009) showed that mental health professionals consider interventions that target the most relevant cause of a mental disorder more effective than interventions that target weaker causes. Research also indicates that participants prefer interventions, which target the first event within a causal chain (Flores, Cobos, Lopez, & Godoy, 2014; Flores, Cobos, López, Godoy, & González-Martín, 2014; Yopchik & Kim, 2009, see Ahn, Kim, & Lebowitz, this volume). The findings of Flores and colleagues (2014) indicate that this preference is probably due to people's notion that by addressing the first cause, all variables within the chain could be changed in desirable way.

## 4. Psychological theories of causal decision making

Many psychological theories of decision making allow for causal beliefs to affect decisions, because these beliefs may influence the subjective likelihood of outcomes (see Box 1 for an overview of respective findings). For example, Kahneman (2011) claims that causal beliefs bias subjective likelihood, which in turn affects decisions (see also Tversky & Kahneman, 1980). We focus more narrowly on three theories in which causal reasoning plays a central role: The causal model theory of choice (Sloman & Hagmayer, 2006), the story model of decision making (Pennington & Hastie, 1992), and attribution theory (Weiner, 1985).

### 4.1 Causal models and the causal model theory of choice

The *Causal Model theory of Choice* (Sloman & Hagmayer, 2006; Hagmayer & Sloman, 2009) provides a descriptive model of decision making based on causal models. A causal model is a mental representation of the causal structure of some set of objects or events in the world (Waldmann, 1996, Sloman, 2005). Causal models contain information about the direction and strength of causal relations connecting the entities. They allow a reasoner to predict the consequences of potential actions and to diagnose the likely cause of an observed outcome (Fernbach, Darlow & Sloman, 2011; Sloman & Lagnado, 2005; Waldmann & Hagmayer, 2005). Hence, causal models can support decision making in the ways predicted by causal decision theory. [2]

According to the causal model theory of choice, decision makers go through three phases. First, the decision maker represents the decision problem as a causal model. This model captures the relevant outcome variables, their potential causes, and the directions and strengths of the causal relations. For example, to decide on a treatment for depression a causal model of the patient's problems (i.e., outcomes) and their potential causes will be constructed. Some potential causes may be observed (e.g., loss of a loved one) others may be inferred from the observed symptoms using the

[2] Causal models and the causal model theory of choice can be formalized using causal Bayes nets (Sloman & Hagmayer, 2006, see Rottmann, this volume for an introduction). A respective causal Bayes net allows computing the causal expected utilities of different options.

causal model (e.g., lack of coping skills). In the second phase, the available options (i.e., potential courses of action) are added to the model. Actions may target outcomes directly (e.g., mood enhancing drugs reducing negative emotions) or causes of outcomes (e.g., training for coping skills). The theory assumes that decision makers conceive of their choice as an intervention. Therefore, the actions to choose from are considered independent of all other variables in the model apart from their direct effects. Once the actions have been added, the model is used to infer the consequences that result from the different options. Inferences are made by using the causal model to run simulations. These simulations take into account uncertainties, including uncertainties with respect to the causal relations. For example, in the case of depression treatment, the consequences of various treatments and treatment combinations are envisioned taking into account that a treatment may not work for a particular patient. In the final phase, a decision is made based on the results of the simulations. If an option has no impact on the desired outcome (i.e., the outcome does not change regardless of whether the option is implemented or not) or its costs outweigh its benefits, it is discarded straight away. If there is only one option that increases the likelihood of the desired outcome, the option is chosen. If several options make the desired outcome more likely, the theory proposes that decision makers adapt their decision making strategy to the given circumstance. If there is time pressure, they may prefer to take the first option that increases the likelihood of desired consequences to a sufficient degree. If there is enough time and the consequences of the decision are important, the decision maker may prefer to search for the option that maximizes causal expected utility. In the case of depression, one clinician may immediately decide on drug treatment because it directly improves the patient's mood, while another clinician may prefer psychotherapy to target deficits in coping which enable negative events to cause depression. Evidence for the causal model theory of choice comes from studies already described in the previous section (Hagmayer & Sloman, 2009; Robinson et al., 2010).

## 4.2 Causal Narratives and the Story Model of Decision Making

Causal models provide a generic framework for representing causal structure. They can be instantiated in various ways to represent a large set of actually observed and possible cases and to make inferences for many different conditions. For example, the same basic structure can be used represent a case in which depression is due to losing a loved one and a case in which depression is due to learned helplessness.

Causal narratives, by contrast, represent a particular, complex causal chain of events (Pennington & Hastie, 1991; 1992). A narrative represents a sequence of events as it probably happened regardless of whether the events in the chain were typical or not. Its primary function is to make sense of an observed sequence of events and to explain how the events are causally related to each other. Narratives can be used to inform decision-making. First, a narrative allows decision makers to target the causes of a current situation with their actions. For example, if the narrative indicates that a patient's depressive symptoms have been caused by a low level of thyroid hormones, hormone therapy is a reasonable treatment option. Second, a certain type of narrative may be linked with a certain type of decision. For example, in medical decision making, the narratives put forward by patients can be matched to illness scripts, which imply certain diagnoses and treatments (Charlin, Boshuizen, Custers, & Feltovich, 2007).

The most prominent theoretical model of decision making based on causal narratives is the "story-model" (Pennington & Hastie, 1991; 1992). The story model has been developed to account for juror decision making in legal cases, but can also be used to model decision making in other areas. The theory proposes that decision makers spontaneously start to construct a narrative when presented with information on a case. By constructing a narrative, the temporal and causal order of the actual events is reconstructed. Events include actions (e.g., stabbing), physical events (e.g., hemorrhage), and mental states (e.g., hate) that are linked through physical and/or intentional/mental causation. Causal links are generally considered necessary and sufficient. When constructing a narrative given information, knowledge about similar events in the past, and generic knowledge about the world are integrated. For example, decision makers may use their knowledge that hatred causes aggressive behavior. The resulting narrative is a complex causal chain with many side arms contributing to the main causal sequence of events. A narrative explains what happened and why it happened. Therefore, narrative-based decision making is also called explanation-based decision making in the literature (e.g., Pennington & Hastie, 1993). The acceptance of a narrative depends on how well it accounts for all case-specific evidence ("coverage"), the absence of alternative narratives ("uniqueness") and its "coherence" or logical consistency, completeness, and plausibility. If there is only one story that has a high coverage and coherence, the person should accept the narrative and be highly confident that the narrative represents what actually happened.

In order to derive a decision from a narrative, decision makers have to know about options and how they relate to different types of narratives. In legal decision making, jurors get this information from a judge who explains how different types of causal chains translate into verdicts. For example, if a person's intention to kill is the cause of another person's death, it is murder. If there was no prior intention to kill, it is manslaughter. In medical decision making, narratives can be matched to clinical guidelines that provide recommendations for treatment. Narrative-based decision making consists of classifying the constructed narrative as being of a particular type. The type of narrative in turn suggests which actions should be taken. Depending on how well the constructed narrative matches the different types of narratives, decisions become more or less difficult. For example, in the medical field a patient's narrative may match the development of different diseases. The simple narrative Stress $\rightarrow$ Poor diet $\rightarrow$ Gastrointestinal problems may point towards depression, psychosomatic disorder, ulcer, or dietary deficiency. In this case the narrative does not allow for a decision on treatment, but it guides decisions on further examinations.

The story model makes three critical predictions. First, it claims that observations or given evidence is spontaneously organized into a causal narrative. Evidence suggests this is the case. Pennington & Hastie (1986) asked participants to think-aloud while reading through the evidence in a legal case. The thought-protocols showed that participants spontaneously reorganized the evidence into a narrative (cf. Kintsch, 1988).

Second, decisions should be based on the constructed narrative and not purely on the given information. This implies that the same evidence may result in different decisions when different narratives are created. Evidence for this prediction comes from studies showing that participants, who generate different narratives for the same evidence reach different verdicts in legal cases (Pennington & Hastie, 1986; 1988; 1992; Lagnado & Gerstenberg, this volume). For example, Pennigton and Hastie (1988) manipulated the order in which evidence was presented to mock jurors. When the order allowed participants to easily construct a story for the defense but not for the prosecution only minority found the defendant guilty. When the same information was re-ordered so that it became easy to construct a narrative for the prosecution, but not for the defense, a majority found the defendant guilty.

Finally, the model predicts that a person´s decision and confidence in the decision depend on coverage, coherence, and uniqueness of the constructed narrative as well as the narrative's fit with the available options. Several experiments have shown that coherence of the evidence predicts decisions (Pennington & Hastie, 1988; 1992). Pennington and Hastie (1992) also manipulated the completeness of the presented information. Participants were either given explicit information about the causal relations among events or not. Importantly, the respective causal relations could easily be inferred from the presented events and everyday knowledge. Thus, participants in either condition should be able to construct the same narratives. Nevertheless, participants favored the verdict associated with the narrative for which information about causal links was explicitly provided. Tenney, Cleary, and Spellman (2009) investigated the influence of uniqueness. They showed that participants were highly sensitive to whether an alternative causal narrative could be constructed for the same evidence. In all conditions the evidence incriminated the defendant in a murder case. While in one condition, no alternative explanation was provided, it was pointed out in another set of conditions that other people would have had the possibility, motive, and/or had no alibi for the time of the murder. Guilty verdicts dropped substantially when other possible suspects were pointed out.

Causal narratives have also been investigated in medical decision-making. This research has focused on whether providing decision makers with a causal narrative affects decisions. A systematic review by Winterbottom, Bekker, Conner, and Montgomery (2008) found that narratives influenced decisions more than statistical or general medical information. Participants were more likely to act in the same way as the person in the narrative. The effect of first person narratives tended to be stronger than third person narratives. This is what would be expected if participants used the given narrative to construct a narrative for themselves. However, an influence of narratives was found only in about a third of the studies. Winterbottom and colleagues (2008) also pointed out that at present it is unclear whether such an influence results in better or worse decisions.

Taken together, the evidence shows that causal narratives influence decisions. They seem to be used spontaneously to organize given information into coherent causal chains, which in turn influence decision making.

## 4.3 Causal attribution and attribution theory

The construction of a causal narrative is not the only way by which an explanation can be provided. Often it may suffice to figure out the cause or the set of causes that led to a particular event or situation. These inferences about causes have been called causal attribution (Kelley, 1972; Kelley & Michaela, 1980). Although there are usually many possible causes that may have contributed to the presence of a target event, people seem to have the tendency to select only one or a few factors as the cause (Weiner, 1985; Kelley, 1972).

*Attribution Theory* (Weiner, 1985; 1986) describes how people make causal attributions and how these attributions affect emotion, motivation, and subsequent behavior. Attributions can be classified with respect to three dimensions: locus of causality (internal vs. external), stability (stable vs. unstable), and controllability (controllable vs. uncontrollable). Thus potential causes can be grouped into eight possible categories. Table 2 shows an example for a classification of potential causes of illness (cf. Roesch & Weiner, 2001). It is important to note that the same causal factor may be classified differently depending on the specific circumstances under which the attribution is made. For example, skills are stable and cannot be modified (i.e., controlled) in the short run, but they can be changed in the long run. The same is true for addictive behaviors, which can be controlled if enough time and external support is provided.

>> Insert Table 2 about here <<

According to attribution theory these categories are meaningful because the respective attributions suggest different emotional, motivational, and behavioral reactions. More precisely, these categories of causes are assumed to be better predictors of emotion, motivation, and action than the individual causes themselves (Weiner, 1985). Stability of the attributed cause should be linked to expected outcomes, which in turn should influence the motivation for action. If the cause is considered stable, the same event or outcome will be expected in the future with increased certainty. By contrast, if the cause is perceived to be unstable, then expectancies should remain uncertain or another outcome may be expected on future occasions. For example, attributing gastrointestinal problems to stress at work implies that the problems will go away when the stressful situation passes. Assumptions about controllability should also affect subsequent actions. If a negative event like illness is attributed to an uncontrollable cause, hopelessness should result, which in turn should decrease the tendency to act (cf. Seligman, 1972). By contrast, an assumption of control should result in a willingness to act. Finally, the assumed locus of causality is important, because it implies whether changing one’s own behavior or intervening on the environment should be preferred, given that there is control over the respective cause. Hence all three dimensions should affect decision making and the motivation to implement any decision made. For example, when cardio-vascular problems are attributed to adopted lifestyle (internal, stable, controllable), then the motivation should rise to change the lifestyle choices, which in turn should result in behavioral changes. By contrast, when the same problem is attributed to hereditary factors (internal, stable, uncontrollable) then there should be little inclination to change one’s lifestyle.

Although having a much broader scope including emotion and motivation, attribution theory can be used as a model of decision making. The process of decision making can be described as follows. First the person encounters an unexpected event or problem that prompts a causal explanation and requires the person to decide how to proceed. Second, the event or problem is attributed to a set of causes. Observed information and knowledge retrieved from memory provide the basis for these inferences. The inferred causes are classified as internal or external, stable or unstable, and controllable or uncontrollable. Depending on the attribution, expectations about the future state of the causes and the effectiveness of potential actions change. Only actions that target controllable causal factors should be judged effective. Therefore, actions controlling stable factors contributing to a problem should be preferred over actions that try to address factors that are uncontrollable or would change on their own without any intervention on behalf of the decision maker.

There is a lot of evidence showing that causal attribution affects decision making. Roesch and Weiner (2001) looked at causal attribution in the medical domain and reviewed studies that investigated the relation between causal attribution, coping behavior and psychological adjustment in patients suffering from illness. All three dimensions of attribution theory were related to coping decisions made by patients. An internal locus of causality was related to more problem-focused coping, assumed control was related to actively approaching rather than avoiding the condition, as was stability. Stability also predicted more problem-focused coping. Moreover, patients that assumed control were better adjusted to their condition than patients who did not. These findings are consistent with findings from studies based on the self-regulatory model of illness (Leventhal, Diefenbach, & Leventhal, 1996), which also found that assumed control was related to being more active in dealing with illness and, in consequence, higher wellbeing and better recovery (cf. Lobban et al., 2003).

Gurevitch, Kliger, and Weiner (2011) investigated the impact of causal attribution on economic decisions (see also Gurevitch & Kliger, 2013). They asked participants to allocate money won in a trivia game to themselves and their partner. Participants were presented with eight different scenarios, which corresponded to the eight categories of causes generated by the three dimensions. For example, in one condition they were told that their teammate got lucky and received easier questions than the others (external, uncontrollable, unstable cause), in another they were told that they themselves had better ability than the others (internal, uncontrollable, stable cause). As expected causal attributions had a substantial effect on how the price money was divided up. Unsurprisingly locus had the strongest effect with the person being responsible for winning the game receiving more. Control also had an effect. If a person had control over the causally responsible factor she received more. Stability did not have an effect, which makes sense as expectancies about the future were irrelevant for the decision to be made.

In the context of economic decision making, Onifade, Harrison, and Cafferty (1997) explored whether causal attribution was related to escalation of commitment (Staw, 1981). They investigated whether causal attribution predicted participants' decision to continue funding a poorly performing project. They found that the assumed stability of the causes for the poor performance was the best predictor for the continuation of the project. When participants got information that the causes were unstable (i.e. due to luck or lack of effort) they were more willing to provide additional funding. Locus of causality only had a minor effect. If causes were assumed to be internal the tendency to continue the project was higher. Further analyses showed that stability and locus affected decisions by changing the expectancy of future success, as predicted by attribution theory.

In sum, the evidence suggests that causal attribution affects decision making, as specified by attribution theory. Which of the causally relevant dimensions (locus, stability, and control) is most relevant depends on the specific decision problem. In general, control seems to be most important, because if there is no control, any action would be futile. Stability is also important because it determines whether taking action is required at all. Unstable causes may resolve themselves without intervention.

4. Discussion

4.1 Summary

Causal decision theory shows that rational decision makers should take the causal structure underlying a decision problem into account, infer the causal consequences of choosing the available options, and maximize causal rather than evidential expected utility when making a choice. Research on decision making has found that decision makers tend to behave in line with these principles.

From a psychological perspective, causal decision theory requires decision makers to engage causal reasoning during decision making. Most theories of decision making, which assume that decisions are based on expected outcomes, do not assume that decision makers use causal reasoning although they would concede that causal beliefs may affect decisions by altering expectancies of outcomes (see Box 1 for evidence). This is true for subjective expected utility theory (Savage, 1954; von Neumann & Morgenstern, 1947), which still is the gold standard for rational decision making, and prospect theory (Kahneman & Tversky, 1979), which is the most influential descriptive theory (cf. Kahneman, 2011). Theories which claim that decisions are based on cue-based heuristics (e.g., Gigerenzer & Goldstein, 1996), may assume that causal beliefs affect the search for and the weighing of cues (e.g., Garcia-Retamnero & Hoffrage, 2006). Finally, theories, which suggest that decisions are guided by scripts (e.g., Abelson, 1981; Schank & Abelson, 1977), assume that scripts include beliefs about typical causal sequences of events and action rules specifying what to do when a specific event sequence unfolds.

We focused on three psychological theories, which assume a central role for causal reasoning in decision making: The causal model theory of choice (Sloman & Hagmayer, 2006; see section 4.1), the story model of decision-making (Pennington & Hastie, 1992, see Section 4.2), and attribution theory (Weiner, 1985; see section 4.3). There are some commonalities and differences between these theories, both in general and with respect to decision making. All three theories assume that people consider the causal structure underlying a decision problem, but they disagree about its representation. Attribution theory assumes that it is represented as a simple model with one or a few causes generating the observed situation. Theories of narrative-based decision making assume that a rather long and complex causal chain is constructed, which describes how the situation developed over time. The causal model theory of choice assumes that a more or less complex causal model is construed, which does not directly represent the temporal development over time. All three theories agree that causal background knowledge is involved in the construction process, but there seem to be differences in the degree to which the resulting model is specific for the individual case. Causal model theory conceptualizes a causal model of a particular problem as an instantiation of a generic model for the type of problem. Observed or given case-specific information is used to instantiate the model for the specific case. The story model and attribution theory assume that an individual, specific model or narrative is construed for the particular case at hand, which is informed by generic knowledge and specifics of the case. The narrative or the attributed causes may deviate strongly from a generic model of the type of problem. The three theories also make different assumptions about how a decision is reached. The causal model theory of choice assumes that decisions are based on the expected causal consequences derived from the causal model. The story model assumes that the constructed narrative is compared to other narratives, which are linked to decisions. Hence, the decision is based on a match between narratives. Attribution theory assumes that causal attributions affect expectations about future states and the effectiveness of actions, which in turn influences decisions.

Predictions of all three theories have support in the literature. However, decision-making does not always follow the predictions of causal theories. Decision makers do not always consider the causal structure underlying a decision problem, nor do they always base their decisions on the causal consequences that would result from their choice (see section 3.1 and 3.2). In the case of self-deception and self-handicapping decision makers choose in a way that does not maximize causal expected utilities. Moreover, decisions can be dominated by considerations independent of causality, such as moral rules (Liu & Ditto, 2013) and social norms (Ajzen, 1991). When decision makers are faced with complex, dynamic systems, they may be unable to use causal reasoning effectively (Sterman, 2000; Osman, 2010; Dörner, 1997). Finally, as we pointed out in the introduction, there might be good reasons not to engage in causal analysis. If the decision maker’s causal background knowledge is very rough or likely to be wrong, then it might be better not to consider causal knowledge. For example, a diabetic who does not recognize the causal influence of his lifestyle choices is likely to end up in poorer health than a diabetic, who just adheres to his doctor's advice and ignores his own causal beliefs (Barnes, Moss-Morris, & Kaufusi, 2004). These findings show that the processes proposed in causal theories are only part of a rich set of strategies that people use to make decisions.

### 4.2 Open Questions

One challenge will be to integrate the different accounts into a more complete theory of causal decision making. Each of the three theories discussed here focuses on a particular type of causal reasoning. Attribution theory focuses on the process of inferring the cause or causes of an event. Narrative theories focus on the process of inferring the complex sequence of events that preceded and generated the decision situation. Causal model theory focuses on generic causal models. It describes how generic models are instantiated for a specific case, how inferences are drawn with respect to unobserved variables, and how consequences of potential actions are inferred. Causal model theory (Sloman, 2005; Waldmann, 1996) seems to be the most promising candidate for an integrative account. Attribution can be conceptualized as a form of diagnostic reasoning based on a causal model (Meder, Mayrhofer & Waldmann, 2014; Fernbach, Darlow and Sloman, 2010, 2011). Thus attribution may be conceptualized as a step in the construction of a causal model of the decision problem. A causal narrative could be described as a complex causal model (Fenton, Neil, & Lagnado, 2013). By assigning a temporal index to the variables, the causal and temporal sequence could be represented within a causal model. The construction of a causal narrative might be an alternative to the construction of a structural causal model if the development of the situation over time is of interest. Thus, an integrative account of causal decision making based on causal models should be possible.

Another challenge is to account for obvious violations of the causal logic of choice, as in self-deception. Bodner and Prelec (2003, Mijovic-Prelec & Prelec, 2010) suggested that people infer the causal expected utility of an action as well as its *diagnostic utility*. The diagnostic utility of an action arises because an action may signify (but not cause) features that have a particular value for the person. Participants in Quattrone and Tversky's experiment (1984) change their tolerance for cold water, because they assume tolerance to indicate a strong heart. Fernbach, Sloman, and Hagmayer (2013, Sloman et al., 2011) used causal models and the idea of diagnostic utility to provide an account for self-deception. They propose that self-deceptive behavior arises when people are unsure about the causes of their actions. Quattrone and Tversky's subjects had uncertainty about the extent to which their endurance was determined by their true pain tolerance versus their desire to achieve a particular result. This uncertainty allows decision makers to choose an action (e.g. high endurance), but to treat the action as diagnostic of a desired trait.

The third challenge will be to clarify the limits of causal decision making. Causal decision theory recommends that decision makers should analyze the causal structure underlying a decision problem. When the same type of decision is made repeatedly, however, a causal analysis may not be necessary. The decision maker may simply learn what the best option is (instrumental learning, e.g., Colwill & Rescorla 1990) and what the best way to make a decision is (learning of decision strategies, e.g., Rieskamp & Otto, 2006). Danks (2014) provides a rational analysis of learning in decision making and shows that causal learning should only ensue when the resulting causal knowledge enables better decisions in the future. This is the case when new options may become available, whose consequences could be predicted from acquired causal knowledge but not based on instrumental learning. Currently very little is known about whether people use causal learning adaptively when making repeated choices. Research by Steyvers, Tenenbaum, Wagenmakers, and Blum (2003), Hagmayer and Meder (2013), and Coenen and colleagues (2015) provide some evidence that people may do so (but see Fernbach & Sloman 2009). We still have a long way to go before we fully understand the role of causality in decision making, its interaction with learning, and – maybe most importantly – its limits.

5. References


Abelson, R.P. (1981). Psychological status of the script concept. *American Psychologist, 36* (7), 715-729.

Ahn, W.-K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*, 147–182.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211.

Ariely, D. & Norton, M.I. (2008). How actions create—not just reveal—preferences. *Trends in Cognitive Sciences*, *12*, 13–16.

Baines, T. & Wittkowski, A. (2013). A systematic review of the literature exploring illness perceptions in mental health utilising the self-regulation model. *Journal of Clinical Psychology in Medical Settings*, 20, 263-274.

Bargh, J.A., & Chartrand, T.L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

Barnes, L., Moss-Morris, R., & Kaufusi, M. (2004). Illness beliefs and adherence in diabetes mellitus: a comparison between Tongan and European patients. *New Zealand Medical Journal*, *117*, 743-751.

Bodner, R., & Prelec, D. (2003). The diagnostic value of actions in a self-signaling model. *The Psychology of Economic Decisions*, *1*, 105–123.

Botti, S., & McGill, A. (2011). The locus of choice: personal causality and satisfaction with hedonic and utilitarian decisions. *Journal of Consumer Research*, *37*, 1065-1978.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid diagnostic signs. *Journal of Abnormal Psychology*, *74*, 271-280.

Charlin, B., Boshuizen, H.P., Custers, E.J., & Feltovich, P.J. (2007). Scripts and clinical reasoning. *Medical Education, 41,* 1178-1184.

Coenen, A., Rehder, B., & Gurekis, T. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102-133.

Colwill, R.M., & Rescorla, R.A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning and Behavior*, *18*, 71-82.

Danks, D. (2014). *Unifying the mind. Cognitive representations as graphical models*. Cambridge, MA: MIT Press.

DeCharms, R. (1968). Personal causation: The internal affective deter- minants of behavior. New York:Academic Press.

DeKwaadsteniet, L., Hagmayer, Y., Krol, N., & Wittman, C. (2010). Causal client models in selecting effective interventions: A cognitive mapping study. *Psychological Assessment, 22*, 581-592.

Doerner, D. (1997). The logic of failure: recognizing and avoiding error in complex situations. Cambridge, MA: Perseus Books.

Fenton, N., Neil, M., & Lagnado, D.A (2013). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive Science*, *37*, 61–102.

Fernbach, P. M., Darlow, A. & Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychological Science, 21,* 329-336.

Fernbach, P. M., Darlow, A. & Sloman, S. A. (2011). When good evidence goes bad: The weak evidence effect in judgment and decision-making. *Cognition*, *119*, 459-467.

Fernbach, P. M. & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, *35*, 678-693.

Fernbach, P.M,  Sloman, S.A., & Hagmayer, Y. (2014). Effort denial in self-deception. *Organizational Behavior and Human Decision Processes*, *123*, 1-8.

Fernbach, P. M., Rogers, T., Fox, C. R. & Sloman, S. A. (2013). Political extremism is supported by an illusion of understanding. *Psychological Science*, 24 (6), 939-946.

Fisher, R. A. (1958). Lung cancer and cigarettes. *Nature*, *182*, 108.

Flores, A., Cobos, P. L., López, F. J., & Godoy, A. (2014).  Detecting fast, on-line reasoning processes in clinical decision making. *Psychological Assessment*, *26*, 660-665.

Flores, A., Cobos, P. L., López, F. J., Godoy, A., & González-Martín, E. (2014). The influence of causal connections between symptoms on the diagnosis of mental disorders: Evidence from on-line and off-line measures. *Journal of Experimental Psychology: Applied*, 20, 175-190.

Furnham, A. (1988). *Lay theories. Everyday understanding of problems in the social sciences*. Oxford: Pergamon Press.

Garcia-Retamero, R. & Hoffrage, U. (2006). How causal knowledge simplifies decision-making. *Minds and Machines*, *16*, 365-380.

Garcia-Retamero, R., Wallin, A., & Dieckmann, A. (2007). When one cue is not enough: Combining fast and frugal heuristics with compound cue processing. *Quarterly Journal of Experimental Psychology*, *60*, 1197–1215.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review, 103, 650-669.

Gurevitch, G., Kliger, D., & Weiner, B. (2011). The role of attribution of causality in economic decision making. *The Journal of Socio-Economics*, *41,* 439-444.

Gurevitch, G, & Kliger, D. (2013). The Manipulation: Socio-economic decision making. *Journal of Economic Psychology*, *39*, 171-184.

Hagmayer, Y. & Meder, B. (2013). Repeated causal decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *39*, 33-50.

Hagmayer, Y., Meder, B., Osman, M., Mangold, S., & Lagnado, D. (2010). Spontaneous causal learning while controlling a dynamic system. *The Open Psychology Journal*, *3*, 145–162. Retrieved from http://www.benthamscience.com/open/topsyj/articles/V003/SI0088TOPSYJ/145TOPSYJ.pdf

Hagmayer, Y., & Sloman, S. A. (2009). People conceive of their choices as intervention. *Journal of Experimental Psychology: General*, *138*, 22–38.

Howard, C.E., & Porzelius, L.K. (1999). The role of dieting in binge eating disorder: Etiology and treatment implications. *Clinical Psychology Review*, *19*, 25-44.

Huber, O., Huber, O.W., & Bär, A. S.(2011). Information Search and mental representation in risky decision making: The advantages first principle. *Journal of Behavioral Decision Making*, *24*, 223-248.

Huber, O., Wider, R., & Huber, O.W. (1997). Active information search and complete information presentation in naturalistic decision tasks. *Acta Psychologica*, *95*, 15-29.

Joyce, J. M. (1999). *The foundations of causal decision theory*. Cambridge, England: Cambridge University Press.

Kahneman, D. (2011). *Thinking fast and slow*. London: Penguin Books.

Kahneman, D., & Miller, D.T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review,* 93, 136-153.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: an analysis of decision under risk. *Econometrica*, *47*, 263-291.

Kahneman, A., & Tversky, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453-458.

Keil, F. C. (2008). Explanation and understanding. *Annual Review of Psychology*, *51*, 227–254.

Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Knaouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution, Perceiving the causes of behavior. Morristown, NJ: General Learning Press.

Kelley, H. H., & Michaela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457-501.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95 (2)*, 163-182.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge: MIT Press.

Koehler, D., & Harvey, N. (2007). *Blackwell Handbook on Decision Making*. Oxford: Blackwell Publishing.

Koslowski, B. (1996). Theory and evidence: The development of scientific reasoning. Cambridge, MA: MIT Press.

Lawson, R. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory & cognition*, 34(8), 1667-1675.

Ladouceur, R., Ferland, F., & Fournier, P. (2003). Correction of erroneous perceptions among primary school students regarding the notions of chance and randomness in gambling. *American Journal of Health Education, 34*, 272-277.

Langer, E. J. (1975). The illusion of control*. Journal of Personality and Social Psychology*, *32*, 311–328.

Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology*, *32*, 951–955.

Leventhal, H., Diefenbach, M., & Leventhal, E.A. (1992). Illness cognition: using common sense to understand treatment adherence and affect cognition interactions. *Cognitive Therapy and Research, 16,* 123-163.

Lewis, D. (2006). Causal decision theory. *The Australasian Journal of Philosophy*, *59*, 5-30.

Lipshitz , R., & Klein, G. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Deciosion Making, 14*, 331-352.

Liu B. S., Ditto P. H. (2013). What dilemma? Moral evaluation shapes factual beliefs. *Social Psychological and Personality Sciences*, *4*, 316–323.

Lobban, F., Barrowclough, C., & Jones, S. (2003). A review of the role of illness models in severe mental illness. *Clinical Psychology Review, 23,* 171-196.

Maher, P. (1987). Causality in the logic of decision. *Theory and Decision*, *22*, 155-172.

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*, 277-301.

Mosing, M. A., Madison, G., Pedersen, N.L., Kuja-Kalkola, R. & Ullen, F. (2014). Practice does not make perfect. No causal effect of music practice on music ability. *Psychological Science*, *25*, 1795-1803.

Murdock, G.P. (1980). *Theories of Illness: A World Survey*. Pittsburgh, PA: University of Pittsburgh Press.

Müller, S.M., Garcia-Retamero, R., Cokely, E., & Maldonado, A. (2011). Causal beliefs and empirical evidence. Decision-making processes in two-alternative forced-choice tasks. Experimental Psychology, 58, 324-332.

Müller, S.M., Garcia-Retamero, R., Galesic, M., & Maldonado, A. (2013). The impact of domain-specific beliefs on decisions and causal judgments. *Acta Psychologica*, *144*, 472-480.

Mijovic-Prelec, D., & Prelec, D. (2010). Self-deception as self-signalling: A model and experimental evidence. *Philosophical Transactions of the Royal Society B*, *365*, 227–240.

Nozick, R. (1993). *The nature of rationality*. Princeton: Princeton University Press.

Onifade, E., Harrison, P.D., & Cafferty T.P. (1997). Causal Attributions for Poorly Performing Projects: Their Effect on Project Continuation Decisions. *Journal of Applied Social PSychology, 27*, 439-452.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin*, *136*, 65-86.

Pennington, N., & Hastie, R., (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology, 51* (2), 242-258.

Pennington, N., & Hastie, R., (1988). Explanation-based decision making: effects of memeory structure on judgement. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14* (3), 521-533.

Pennington, N., & Hastie, R. (1991). A cognitive theory of juror decision making. The story model. *Cardozo Law Review*, *13*, 519-558.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology, 62* (2), 189-206.

Pennington, N., & Hastie, R., (1993). Reasoning in explanation-based decision making. *Cognition, 49*, 123-163.

Quattrone, G., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology*, *46*, 237–248.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory on how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.

Robinson, E., Sloman, S.A., Hagmayer, Y., & Hertzog, C. (2010). Causality in solving economic problems. *The Journal of Problem Solving*, *3*, 106-130.

Roesch, S.C., & Weiner, B. (2001). A meta-analytic review of coping with illness. Do causal attributions matter? *Journal of Psychosomatic Research, 50,* 205-219.

Rozenblit, L., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521-562.

Savage, L.J. (1954). *The foundations of statistics*. New York: Wiley.

Schank, R.C., & Abelson, R.P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Seligman, M.E. (1972). Learned helplessness. *Annual Review of Medicine, 23*, 407-412.

Skyrms, B. (1982). Causal decision theory. *The Journal of Philosophy, 79*, 695-711.

Sloman, S. A. (2005) Causal models: How we think about the world and its alternatives. Oxford: Oxford University Press.

Sloman, S. A., & Hagmayer, Y. (2006). The causal psycho-logic of choice. Trends in Cognitive Sciences, 10, 407-412.

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

Sloman, S.A. , Fernbach, P.M., &  Hagmayer Y. (2010). Self deception requires vagueness. Cognition, 115, 268–281.

Staw, B.M., (1981). The Escalation of Commitment To a Course of Action. *Academy of Management, 6,* 577-587.

Sterman, J. D. (2000). *Business dynamics. Systems thinking and modeling for a complex world*. Boston: McGraw-Hill.

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

Stockford, K., Turner, H. & Cooper, M. (2007). Illness perception and its relationship to readiness to change in the eating disorders: a preliminary investigation. *British Journal of Clinical Psychology, 46*, 139-54.

Sylvain, C., Ladouceur, R. & Boisvert, J. M. (1997). Cognitive and behavioural treatment of pathological gambling: a controlled study. *Journal of Consulting and Clinical Psychology, 65,* 727–732.

Tenney, E.R., Cleary, M.D., & Spellman B.A. (2009). Unpacking the doubt in “beyond a reasonable doubt”: plausible alternative stories increase not guilty verdicts. *Basic and Applied Social Psychology, 31,* 1-8.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), Progress in social psychology (pp. 49-72). Hillsdale, NJ: Erlbaum.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and behavior*. Princeton, NJ: Princeton University Press.

Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34: Causal learning (pp. 47-88). San Diego: Academic Press.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing vs. doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning memory and Cognition, 31*, 216-227.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548-573.

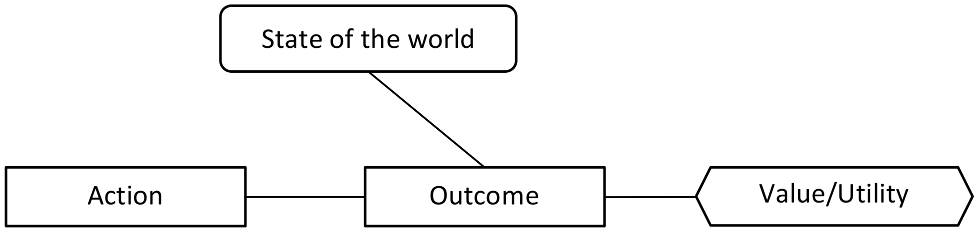
Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.

Winterbottom, A., Bekker, H.L., Conner, M. & Montgomery, J. (2008). Does narrative information bias individual's decision making? A systematic review. *Social Science and Medicine*, *67*, 2079-2088.

Yopchick, J. E., & Kim, N. S. (2009). The influence of causal information on judgments of treatment efficacy. *Memory & Cognition*, *37*, 29-41.

Zsambok, C. E., & Klein, G. (1997). *Naturalistic decision making*. Mahwah, NJ: Erlbaum.

Figure 1: Decision making from the perspective of expected utility theory.
Action A and the state of the world S jointly determine the probability of potential outcomes O. Outcomes have a certain utility for the decision maker.



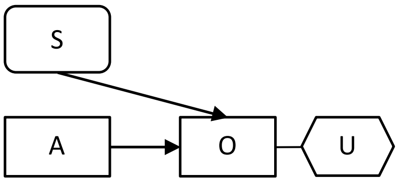

Example decision on breast cancer screening for 40year old women:
Table of actions, states, expectancies, outcomes and resulting expected utilities

| States | Breast Cancer | | | | No Breast Cancer | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Death from Cancer | | Live long | | Death from Cancer | | Live long | | |
| Options | Prob. | Utility | Prob. | Utility | Prob. | Utility | Prob. | Utility | Expected utility |
| Action 1: Screening | .024 | 0 | .096 | 90 | 0 | 0 | .88 | 90 | $EU(A_1)$ = .024*0 + .096*90+0*0+.88*90 = 87.8 |
| Action 2: No screening | .03 | 1 | .09 | 91 | 0 | 1 | .88 | 91 | $EU(A_2)$ = .03*1 + .09*91+0*1+.88*91 = 88.4 |

Figure 2: Decision making from the perspective of causal decision making theories.
Action A, the state of the world S, and the outcome O are causally related to each other. They may be related through different causal structures. Note that all structures (a) – (d) entail that there is a probabilistic relation between action and outcome, that is, the outcome is probabilistically dependent on the action.

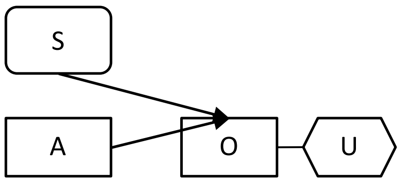
Causal Structures | Calculation of Causally Expected Utility (CEU)

(a) Action generates Outcome



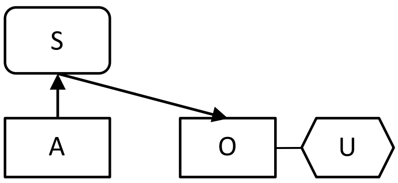

$CEU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$
$= \sum_{ji} P(O_{ji} | S_j, A_i)*P(S_j) * U(O_{ji})$
If the action and the state of the world independently affect the outcome and are independent of each other, then
$= \sum_{ji} [P(O_{ji} | A_i) + P(O_{ji} | S_j) – P(O_{ji} | A_i)* P(O_{ji} | S_j)]*P(S_j) * U(O_{ji})$

(b) Action enables Outcome



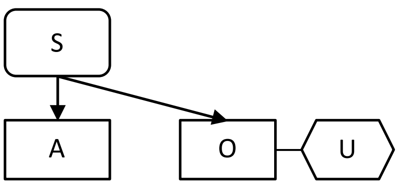

$CEU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$
$= \sum_{ji} P(O_{ji} | S_j, A_i)*P(S_j) * U(O_{ji})$
The action and the state of the world interact in causing the outcome. The action and the state of the world are independent.

(c) Action affects state of the world, which generates outcome




$CEU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$
$= \sum_{ji} P(O_{ji} | S_j)*P(S_j | A_i) * U(O_{ji})$
The state of the world causes the outcome and the state of the world is dependent on the action.

(d) Action and outcome are spuriously related due to state of the world




$CEU(A_i) = \sum_{ji} P(O_{ji}) * U(O_{ji})$
$= \sum_{ji} P(O_{ji} | S_j)*P(S_j) * U(O_{ji})$
The state of the world causes the outcome. The probability of the state of the world is causally independent of the action.

Figure 3: Difference between deliberately choosing an action and observing an action. Solid arrows indicate causal relations, dashed arrows indicate evidential, statistical relations.

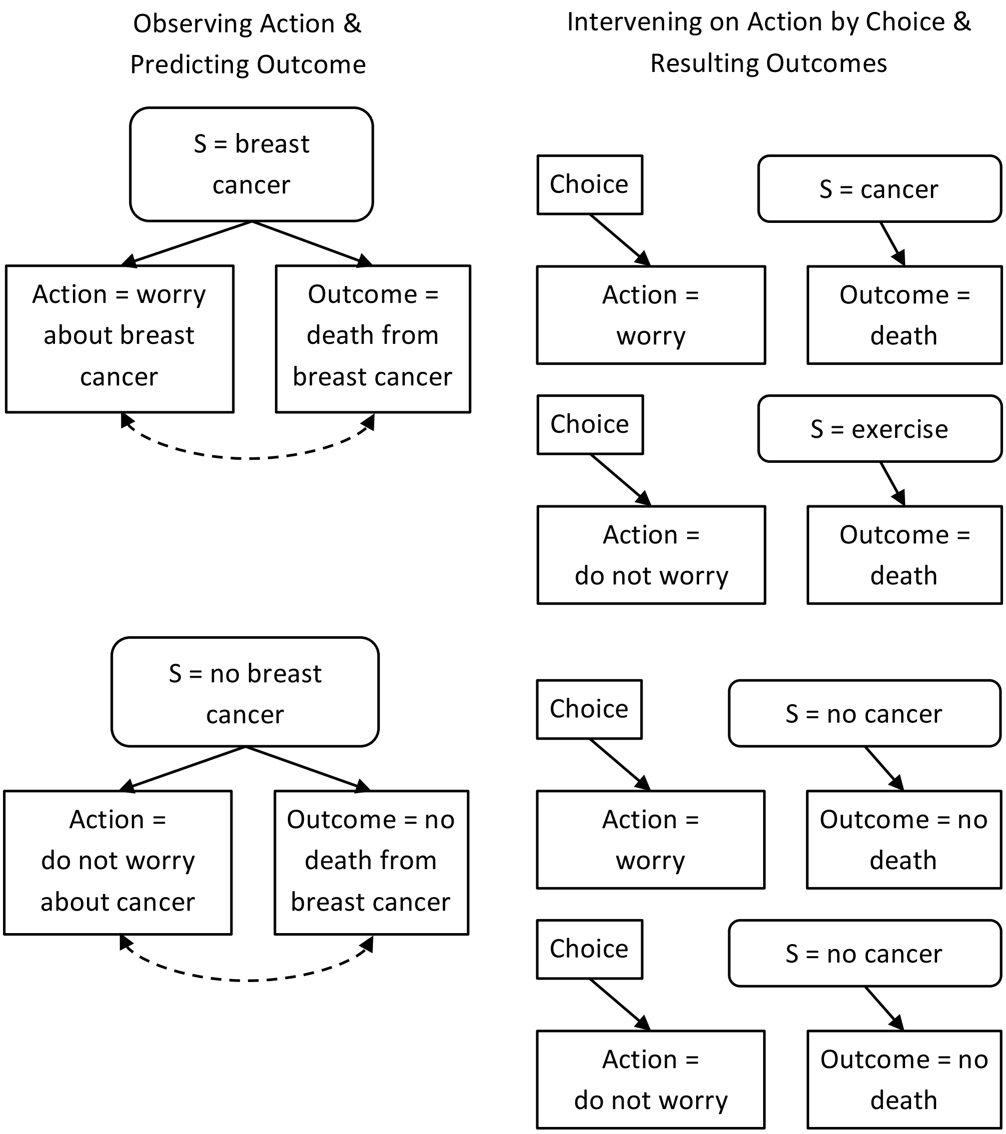


Table 1. *Matrix of options and expected payoffs in Experiment 1 of Robinson, Sloman, Hagmayer, & Hertzog (2010). Expected payoffs of the participant are given in bold print. Participants were also told that they and their competitor made the same decision in the past 90% of the time. Participants had to decide whether to purchase dollars or euros.*

| | Your competitor buys dollars | Your competitor buys euros |
|---|---|---|
| **You buy dollars** | **1 billion $** / 1 billion $ | **50 million $** / 1.2 billion $ |
| **You buy euros** | **1.2 billion $** / 50 million $ | **100 million $** / 100 million $ |

Table 2. *Categories of causal attributions in the medical domain (adapted from Roesch & Weiner, 2001)*.

| Stability / Control / Locus of causality | Stable: Controllable | Stable: Uncontrollable | Unstable: Controllable | Unstable: Uncontrollable |
|---|---|---|---|---|
| Internal | Some physiological processes | Heredity, Personality, Some physiological processes | Own actions, Own effort | |
| External | | Fate[3], Economic environment | Some environmental stressors | Chance, Stimuli controlling behavior Some environmental stressors |

[3] Roesch and Weiner (2001) considered fate to be unstable. However, anthropological and ethnographic research shows that fate is considered to be stable (Murdock, 1980)