

## The interventionist account of causation and the basing relation

Kevin McCain

Published online: 19 February 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** It is commonplace to distinguish between propositional justification (having good reasons for believing  $p$ ) and doxastic justification (believing  $p$  on the basis of those good reasons). One necessary requirement for bridging the gap between  $S$ 's merely having propositional justification that  $p$  and  $S$ 's having doxastic justification that  $p$  is that  $S$  base her belief that  $p$  on her reasons (propositional justification). A plausible suggestion for what it takes for  $S$ 's belief to be based on her reasons is that her reasons must contribute causally to  $S$ 's having that belief. Though this suggestion is plausible, causal accounts of the basing relation that have been proposed have not fared well. In particular, cases involving causal deviancy and cases involving over-determination have posed serious problems for causal accounts of the basing relation. Although previous causal accounts of the basing relation seem to fall before these problems, it is possible to construct an acceptable causal account of the basing relation. That is, it is possible to construct a causal account of the basing relation that not only fits our intuitions about doxastic justification in general, but also is not susceptible to the problems posed by causal deviancy and causal over-determination. The interventionist account of causation provides the tools for constructing such an account. My aim is to make use of the insights of the interventionist account of causation to develop and defend an adequate causal account of the basing relation.

**Keywords** Justification · Basing relation · Causation · Interventionist · Causal deviancy

---

K. McCain (✉)

Department of Philosophy, University of Rochester, Box 270078, Rochester, NY 14627-0078, USA  
e-mail: kmccain@mail.rochester.edu

## 1 Introduction

It is commonplace to distinguish between propositional justification (having good reasons for believing  $p$ ) and doxastic justification (believing  $p$  on the basis of those good reasons).<sup>1</sup> A theory that only provides an account of propositional justification is inadequate because an account of doxastic justification is needed for a satisfactory analysis of knowledge.<sup>2</sup> One of the key requirements for providing an adequate account of doxastic justification is to explain what is required in addition to propositional justification that  $p$  in order for  $S$  to have doxastic justification that  $p$ . One necessary requirement for bridging the gap between  $S$ 's merely having propositional justification that  $p$  and  $S$ 's having doxastic justification that  $p$  is that  $S$  base her belief that  $p$  on her reasons (propositional justification). So, any adequate theory of doxastic justification must provide an explanation of what it takes for  $S$  to base her belief that  $p$  on her reasons.<sup>3</sup>

A plausible suggestion for what it takes for  $S$ 's belief to be based on her reasons is that her reasons must contribute causally to  $S$ 's having that belief. Though this suggestion is plausible, causal accounts of the basing relation that have been proposed have not fared well.<sup>4</sup> In particular, cases involving causal deviancy and cases involving over-determination have posed serious problems for causal accounts of the basing relation. Although previous causal accounts of the basing relation seem to fall before these problems, it is possible to construct an acceptable causal account of the basing relation. That is, it is possible to construct a causal account of the basing relation that not only fits our intuitions about doxastic justification in general, but also is not susceptible to the problems posed by causal deviancy and causal over-determination. The interventionist account of causation provides the tools for constructing such an account. My aim is to make use of the insights of the interventionist account of causation to develop and defend an adequate causal account of the basing relation.

## 2 Interventionist account of causation

Before constructing the interventionist account of the basing relation, it is important to be clear about some of the general features of the interventionist account of

<sup>1</sup> For expressions of this distinction see Audi (1983), Bergmann (2006), Comesana (2006, 2010), Feldman and Conee (1985), and Pollock and Cruz (1999).

<sup>2</sup> Granted some, such as Williamson (2000), deny that there can be any satisfactory analysis of knowledge in terms of its components. However, insofar as we do seek an analysis of knowledge, an account of doxastic justification will be necessary. And even if there is no analysis of knowledge, it is still plausible to think that doxastically justified belief is a *necessary* condition for knowledge.

<sup>3</sup> Comesana (2006) and (2010) argues that any account of doxastic justification must utilize the basing relation. Bergmann (2006) disagrees. He argues that basing is not necessary for doxastic justification because propositional justification is not necessary for doxastic justification. Although it is not clear that Bergmann's argument is successful; even if we assume that Bergmann is correct, any acceptable account of doxastic justification must still provide an explanation of the basing relation in order to account for inferentially justified beliefs.

<sup>4</sup> See Korcz (1997) for a survey of various attempts to provide an account of the basing relation and problems with those accounts.

causation. The interventionist account of causation is a kind of manipulability theory of causation. In general manipulability theories of causation claim that when *A* is a cause of *B* our manipulating *A* in certain ways should be a way for us to manipulate, or change, *B*. More precisely, the primary intuition at work in manipulability theories can be formulated in the following manner:

*A* causes *B* if and only if *B* would change if an appropriate manipulation on *A* were to be carried out.<sup>5</sup>

The key to providing an acceptable version of a manipulability theory is to satisfactorily define what counts as ‘an appropriate manipulation’. Although manipulability theorists have proposed various strategies for specifying ‘appropriate manipulations’, the most promising strategy is to do so in terms of interventions.<sup>6</sup> In other words, the interventionist account of causation is the most promising version of a manipulability theory of causation.<sup>7</sup> For our purposes it will not be necessary to examine the interventionist account of causation in painstaking detail. Instead, our purposes will be served with a sketch of some of the key features of the account.<sup>8</sup>

The first feature of the interventionist account of causation that we need to examine is the idea of an *intervention*. Interventions are manipulations upon one or more variables in a system under idealized experimental conditions. An intervention on a variable *X* should be understood in terms of experimental manipulations of *X* that are well designed for determining if *X* causes *Y* in an idealized experimental setting (an experimental setting that excludes confounding influences). More precisely, “an intervention *I* on *X* with respect to *Y* will be such that *I* causes a change in *X*, *I* does not cause a change in *Y* via some route that does not go through *X*, and *I* is exogenous in the sense of not itself having a cause that affects *Y* via a route that does not go through *X*.”<sup>9</sup> Further, interventions are “*surgical*, that is, in the sense that the usual causes of the variable in question are suspended, so that the value of the variable depends only on the intervention.”<sup>10</sup> In other words, *I* will be

<sup>5</sup> This is Woodward’s (2008a) formulation. Woodward (2003) expresses the intuition behind manipulation theories with a slogan “No causal difference without a difference in manipulability relations, and no difference in manipulability relations without a causal difference.” p. 61.

<sup>6</sup> See von Wright (1971) and Menzies and Price (1993) for recent attempts to define ‘appropriate manipulation’ by appealing to free actions instead of interventions. See Woodward (2003) for the most complete account of manipulations in terms of interventions yet produced.

<sup>7</sup> Given the focus of this paper, I will not spend time defending this claim. The reader is encouraged to consult Woodward (2003) and Woodward (2008a) for a cogent defense of the superiority of the interventionist account of causation over other manipulability theories.

<sup>8</sup> The sketch of the interventionist account of causation that we will consider draws heavily from James Woodward’s account. Readers seeking a more detailed explanation of the interventionist account of causation should see Woodward (2003) and Woodward (2008a). For the purposes of this paper, I will assume that the interventionist account of causation that Woodward develops is the (at least mostly) correct account of causation. Readers who are skeptical of Woodward’s account of causation can understand the current project as arguing for a conditional conclusion: if the relevant features of the interventionist account of causation are correct, then it is possible to use them to construct an adequate account of the basing relation.

<sup>9</sup> Woodward (2008b, p. 213). See Woodward (2003, p. 98) for a formal definition of ‘intervention’.

<sup>10</sup> Brad Weslake “Exclusion Excluded” (unpublished manuscript, p. 3) (His emphasis).

the sole determiner of the value of  $X$ . Simply put, the idea is that an intervention on  $X$  is some sort of change that an experimenter in an ideal setting can bring about in  $X$  that is such that the method of bringing about that change will directly affect only  $X$  and the method of changing  $X$  will exclusively set the value of  $X$ .

The notion of an intervention as defined above provides the foundation for a simple characterization of what it is for  $X$  to be causally relevant to  $Y$ :

(M)  $X$  causes  $Y$  if and only if were some intervention ... that changes the value of  $X$  to occur,  $Y$  or the probability distribution of  $Y$  would change in some regular, stable way [at] least in some range of background circumstances  $B$ .<sup>11</sup>

There are two points about M that we should make clear. First, M adequately captures our intuition that a cause is a difference maker. Given M, we can see how a cause must make a difference to its effect because in order for  $X$  to be a cause of  $Y$ , changes in the value of  $X$  must be correlated with changes in  $Y$  in a stable way. Second, it should be noted that M is only a characterization of what it is for  $X$  to be causally relevant to  $Y$ . As such, accepting M does not commit us to claiming that the only way to identify causal relations is by conducting experimental interventions on  $X$ .

Another feature of the interventionist account of causation that we need to think about is an extension of M which accounts for a particular kind of type-causal relation, the relation of a *direct cause*. The notion of a direct cause that will be relevant for our purposes is the following:

A necessary and sufficient condition for  $X$  to be a direct cause of  $Y$  with respect to some variable set  $V$  is that there be a possible intervention on  $X$  that will change  $Y$  (or the probability distribution of  $Y$ ) when all other variables in  $V$  besides  $X$  and  $Y$  are held fixed at some value by interventions.<sup>12</sup>

So, for example, if we have a variable set that includes  $X$ ,  $Y$ , and  $Z$ ;  $X$  is a direct cause of  $Y$  if and only if there is a possible intervention on the value of  $X$  that will change the value of  $Y$  when we hold fixed the value of  $Z$ . It would be natural at this point to question how we should determine which variable set to use when evaluating whether  $X$  is a direct cause of  $Y$ . As Woodward notes, our purposes for investigating a particular phenomenon will determine which factors we include in the variable set.<sup>13</sup> In many cases this may be a difficult decision. For example, when considering a phenomenon in biology it may be difficult to decide whether the relevant variable set should be a set of micro-physical variables such as the chemical composition of various molecules and features of genetic traits or if the relevant variable set should be a set of higher-level variables such as features of populations and ecological factors. This sort of decision can be very difficult; however, we should keep in mind two important points. First, variable sets can be of mixed levels. That is, we do not have to exclude lower (higher) level variables simply because our variable set includes higher (lower) level variables. Second, for

<sup>11</sup> Woodward (2008b, p. 213).

<sup>12</sup> Woodward (2003, p. 55).

<sup>13</sup> Woodward (2003, pp. 86–90).

our current endeavor the choice of the appropriate variable set is not a difficult matter. We are concerned with when an agent's belief is based on her reasons. So, the relevant variable set for our purposes will be a set that includes variables that can propositionally justify or defeat a belief. Thus, the relevant variable set for us will be a set of predominately higher-level mental states such as beliefs, perceptual states (hallucinations as well as veridical perceptions), desires, and so on.<sup>14</sup>

Before considering the final feature of the interventionist account of causation that is relevant to our endeavor, the concept of an *actual cause*, it is important to first explicate two further notions that are used to define this concept. The first is the notion of a *directed path*. A directed path from one variable,  $X$ , to another variable,  $Y$ , is a "chain of direct causal relationships...from  $X$  to  $Y$ ."<sup>15</sup> Simply put, there is a directed path from  $X$  to  $Y$  if and only if each variable starting with  $X$  and ending with  $Y$  is a direct cause of the variable that immediately succeeds it. For example, if there is a causal chain from  $X$  to  $R$  to  $S$  to  $Y$ , that chain is a directed path from  $X$  to  $Y$  if and only if  $X$  is a direct cause of  $R$ ,  $R$  is a direct cause of  $S$ , and  $S$  is a direct cause of  $Y$ . The second notion that we need to explicate before examining the meaning of 'actual cause' is the notion of a *redundancy range*. A redundancy range can be defined as follows:

For a path  $P$  from  $X$  to  $Y$  in a causal model, define  $V_1 \dots V_n$  as all variables that are not on  $P$ . Values  $v_1 \dots v_n$  are on the redundancy range for  $V_i$  with respect to  $P$  if no intervention on  $v_1 \dots v_n$  while holding  $X$  fixed would result in a change to the actual value of  $Y$ .<sup>16</sup>

Now that we have explicated the notions of directed path and redundancy range, we can turn our attention to the final feature of the interventionist account of causation that we need to examine; the concept of an *actual cause*. As one might expect, when  $X$  is an actual cause of  $Y$ ,  $X$ 's having its actual value is a cause of  $Y$ 's having its actual value; hence 'actual cause' refers to a relation of token causation. According to James Woodward,  $X = x$  ( $X$ 's taking some particular value) is an actual cause of  $Y = y$  ( $Y$ 's taking some particular value) if and only if both of the following conditions are satisfied:

(AC\*1) The actual value of  $X = x$  and the actual value of  $Y = y$ .

(AC\*2) For each directed path  $P$  from  $X$  to  $Y$ , fix by interventions all direct causes  $Z_i$  of  $Y$  that do not lie along  $P$  at some combinations of values within their redundancy range. Then determine whether, for each path from  $X$  to

<sup>14</sup> This is not to say that the relevant variable set will only include variables of this kind. It is plausible to think that neurological damage, chemical imbalances, etc. can affect basing as well.

<sup>15</sup> Woodward (2003, p. 59).

<sup>16</sup> Weslake "Exclusion Excluded" (unpublished manuscript, p. 6). It bears mentioning that although a variable's actual value is always within its redundancy range, it is possible for non-actual values to be within the variable's redundancy range as well. Further, there can be cases in which whether a value falls within a variable's redundancy range depends on the values of other variables. For instance, there can be cases where the actual value of  $V_1$  is 1 and a value of 0 for  $V_1$  falls within its redundancy range only when  $V_2$  is held fixed at some particular value. So, in this case the value 0 is within the redundancy range for  $V_1$  relative to a particular value of  $V_2$ .

$Y$  and for each possible combination of values for the direct causes  $Z_i$  of  $Y$  that are not on this route and that are in the redundancy range of  $Z_i$ , whether there is an intervention on  $X$  that will change the value of  $Y$ . (AC\*2) is satisfied if the answer to this question is “yes” for at least one route and possible combination of values within the redundancy range of the  $Z_i$ .<sup>17</sup>

The basic idea is that when  $X$  is an actual cause of  $Y$  it is possible to change the actual value of  $Y$  by changing the actual value of  $X$  while holding all other direct causes of  $Y$  that are not part of a chain of direct causes (directed path) leading from  $X$  to  $Y$  fixed at a value within their redundancy range.

At this point, it will be instructive to consider an example to help further illuminate some of the features of the interventionist account of causation. Consider the following: Agent Smith attempts to start a fire by planting an incendiary device in an oxygen filled room. Smith’s incendiary device malfunctions. However, a spark occurs because of a short circuit in the room’s electrical wiring. The spark results in a fire. For simplicity, we can assume that the only relevant factors for the fire ( $F$ ) are the functioning of the incendiary device ( $D$ ), the short-circuiting of the electrical wiring ( $S$ ), the spark ( $K$ ), and the oxygen in the room ( $O$ ). So, the variable set corresponding to this example only contains the following variables:  $F$ ,  $D$ ,  $S$ ,  $K$ , and  $O$ . Additionally, we can assume that in this case the only values for  $F$ ,  $D$ ,  $S$ ,  $K$ , and  $O$  are 1 or 0. That is to say, either the fire occurs or it does not; either the incendiary device functions properly or it does not; either the electrical wiring short circuits or it does not; either there is a spark or there is not; and either the room is filled with oxygen or it has no oxygen. In the case as described the value of  $D$  is 0 because the incendiary device malfunctions and the values of  $S$ ,  $K$ ,  $O$ , and  $F$  are each 1 because the electrical wiring short circuits, there is a spark, the room is filled with oxygen, and the fire occurs.

According to the interventionist account of causation, the short-circuiting of the electrical wiring, the spark, and the oxygen in the room are all actual causes of the fire, but the incendiary device is not. If we hold the variables that are not on the directed path from  $S$  to  $F$ , that is  $D$  and  $O$ , fixed at values in their redundancy range, it is possible to change the actual value of  $F$  by changing the actual value  $S$ . The idea is that if we hold fixed the oxygen in the room and the malfunctioning of the incendiary device, intervening so that the electrical wiring does not short circuit will result in there being no fire. The same holds for  $K$  and  $O$ . That is, if we change the actual value of  $K$  so that there is no spark while holding fixed the electrical wiring’s short-circuiting, the incendiary device’s malfunctioning, and the oxygen level of the room; the fire will not occur. Likewise, if we intervene to change the actual value of  $O$  so that there is no oxygen in the room while holding fixed the malfunctioning of the incendiary device, the electrical wiring’s short-circuiting, and the presence of the spark; the fire will not occur. So,  $S$ ,  $K$ , and  $O$  are all actual causes of  $F$ . However, if we hold  $S$ ,  $K$ , and  $O$  fixed at their actual values (which are the only values in their redundancy range with respect to the path from  $D$  to  $F$ ), changing the actual value of

<sup>17</sup> Woodward (2003, p. 84).

$D$  will not result in a change in the actual value of  $F$ . With these variables held fixed at their actual values the value of  $F$  will be 1, that is to say the fire will occur. Intervening so that the incendiary device works will not change the fact that the fire occurs. Thus,  $D$  is not an actual cause of  $F$  in this case because the incendiary device does not make a difference to whether or not there is a fire.

Interestingly, though  $S$  is an actual cause of  $F$  and  $D$  is not,  $D$  is a direct cause of  $F$  and  $S$  is not. Recall that in order for one variable to be a direct cause of another variable there must be a possible intervention on the first variable that will result in a change in the value of the second when all other variables in the set are held fixed at some value. First,  $D$  is a direct cause of  $F$  because if we hold  $O$ ,  $S$ , and  $K$  held fixed at the respective values 1, 0, and 0, intervening on  $D$  will cause changes in  $F$ . More concretely, if we intervene to ensure that the electrical wiring does not short circuit, there is no spark, but there is oxygen in the room; changing whether the incendiary device functions properly or not will change whether the fire occurs. So, the incendiary device is a direct cause of the fire, even though it is not an actual cause of the fire. Second,  $S$  is not a direct cause of  $F$  because we cannot hold  $D$ ,  $K$  and  $O$  fixed at some values and manipulate the value of  $F$  by changing the value of  $S$ . The reason this is the case is that  $S$ 's causal influence on  $F$  must go through  $K$ . Once the value of  $K$  is held fixed, changes in  $S$  will have no effect on  $F$ . In other words, the only way that the short-circuiting of the electrical wiring can affect whether or not there is a fire in this case is by affecting whether or not there is a spark in the room. Once the presence or absence of the spark is held fixed, manipulating whether the electrical wiring short circuits cannot change whether the fire occurs. So, the short circuiting of the electrical wiring is not a direct cause of the fire, even though it is an actual cause of the fire.

Now that we have the relevant features of the interventionist account of causation in hand, we can turn our attention toward developing and evaluating an interventionist account of the basing relation.

### 3 Interventionist account of the basing relation

#### 3.1 The account

As we observed above, a crucial component for bridging the gap between propositional justification and doxastic justification is the basing relation. It is widely agreed that in order for  $S$ 's belief that  $p$  to be doxastically justified her belief must not only be propositionally justified by her reasons, but she must also base her belief on those reasons. The interventionist account of causation provides us with the conceptual tools to give a satisfactory account of basing in general and to give a satisfactory account of what it means for someone to base her belief on her reasons in particular. To begin, here is the interventionist account of the general basing relation:

**(IB):**  $S$ 's belief that  $p$  at  $t$  is based on  $X$ , if and only if  
at  $t$ :

- (1)  $X$  is a direct cause of  $S$ 's believing that  $p$   
AND
- (2)  $X$  is an actual cause of  $S$ 's believing that  $p$ .

When we are discussing whether  $S$ 's belief that  $p$  has doxastic justification we are interested in whether  $S$ 's belief that  $p$  is based on her reasons (her propositional justification). However, we are interested not only in whether  $S$ 's reasons are part of the basis of her belief that  $p$ , we are interested in whether  $S$ 's reasons bear a causal relation of sufficient strength to her belief. So, we are interested in an enhanced instance of the general basing relation:

**(IB-R):**  $S$ 's belief that  $p$  at  $t$  is based on her reasons,  $R$ , if and only if at  $t$ :

- (1) Each  $r_i \in R$  is a direct cause of  $S$ 's believing that  $p$   
AND
- (2) Each  $r_i \in R$  is an actual cause of  $S$ 's believing that  $p$ .  
AND
- (3) It is not the case that intervening to set the values of all direct causes of  $S$ 's believing that  $p$ , other than the members of  $R$ , to 0 will result in  $S$ 's not believing that  $p$  when every  $r_i \in R$  is held fixed at its actual value.

It is important to clarify various points concerning **IB-R** before turning our attention to exploring how **IB-R** works in a variety of cases.<sup>18</sup> First,  $R$  is a subset of  $S$ 's mental states, which propositionally justifies  $p$ .  $S$ 's belief that  $p$  has the connection to her propositional justification needed for doxastic justification only if there is at least some  $R$  such that  $S$ 's belief that  $p$  is based on  $R$  in the manner described in **IB-R** and there is no more inclusive set of  $S$ 's mental states which fails to propositionally justify  $p$ .

Second, the senses of 'direct cause' and 'actual cause' that are relevant for **IB-R** are those employed in the interventionist account of causation, which were defined in Sect. 2.

Third, when evaluating the causal relations between  $S$ 's reasons and her beliefs we will utilize sets of variables. These variable sets will include variables for  $S$ 's belief, her reasons ( $R$ ), and any other factor that may plausibly be a causal influence on  $S$ 's believing as she does. The factors besides  $R$  and  $S$ 's belief that may be of interest to us are other mental states/mechanisms or external influences. For instance, if we are considering a case of someone who engages in wishful thinking, then we will include a wishful thinking variable in the variable set. Also, if we are considering a case that involves a potential direct external influence on  $S$ 's belief (an external influence that affects  $S$ 's believing a proposition, but does not affect or potentially affect her belief by affecting some other mental state/mechanism) such as  $S$ 's being struck by a ray from Alpha Centauri, then a variable representing the Alpha Centaurian ray will be a member of the variable set.

<sup>18</sup> Throughout the remainder of the paper we will be focused on **IB-R** instead of **IB**.



Another point that needs further explication is the idea of assigning values for the variables mentioned. One might worry about what sort of values a mental state or an external influence can have. The simplest valuation schema that we could impose is a binary system. Each of these variables has a value of either 1 or 0. So, for any perceptual state, S is either in that perceptual state or she is not, for any external influence, the influence is either active or it is not, and so on. Although for simplicity we will tend to follow a binary valuation schema, we should not feel limited to such a coarse-grained schema. One way in which we could adopt a more fine-grained valuation schema is to adopt a schema for these variables that assigns values based on the strength of the mental state/mechanism or influence in question. So, for example, we might want to assign values to perceptual states in terms of their degree of vividness; the more vivid the perceptual state the higher the value. We could develop similar fine-grained schemas for any other variables we need to include in a given variable set.

Finally, **IB-R** (and **IB** as well) describes a relation between a belief and its causal base at a particular time. So, **IB-R** offers an account of basing for both belief formation and belief sustention. This is important because the causal basis of a belief may change over time so that at  $t_1$   $x$  causes S to form the belief that p, but at  $t_2$  S's belief that p is causally sustained by  $y$ . Depending on the nature of  $x$  and  $y$  it is possible that S's belief that p is based on her reasons at one time, but not another. So, **IB-R** allows for the intuitively correct result that someone may form a belief that is not based on good reasons, but later gain good reasons for the belief and continue to hold the belief because of those reasons so that the belief at the later time is based on those reasons.

### 3.2 Applications of **IB-R**

In order to fully appreciate the viability of **IB-R**, it will be instructive to consider how **IB-R** applies to a variety of cases of differing complexity. In all of these cases we will be considering whether S's belief is based on her reasons in the sense required for doxastic justification, so we will be assuming that S's reasons propositionally justify her belief. Additionally, we will assume that all of the variables which we examine can only have values of 0 or 1. For instance, when we are considering S's belief that p we will assume that there are only two relevant options: S believes that p and S does not believe that p.<sup>19</sup> Let us begin with a very simple case, one involving a simple perceptual belief. A simple perceptual belief is one that S has on the basis of only her experiences, that is, the belief is non-inferential.<sup>20</sup> The following is an example of this kind of case (Case 1):

<sup>19</sup> It should be noted that the fact that S does not believe that p is consistent with several states: S's believing that not-p, S's withholding belief that p, and S's having no doxastic attitude toward p.

<sup>20</sup> For our purposes, it will not be necessary to worry about the nature of this experience. That is, it is not necessary for us to determine the nature of the content of this experience (whether it is conceptual, non-conceptual, or perhaps a mixture). Also, for our purposes we will set aside issues of concept possession, which may lead one to think that S cannot have a belief solely on the basis of her experience if she lacks the relevant experiential background.

Sally has a perceptual experience of a red patch and as a result she believes there is a red patch.

Assuming for the sake of illustration that there are no other relevant features of Sally or her environment that may contribute causally to her having the belief that there is a red patch, the relevant variable set  $\mathbf{V}_1$  will only contain two variables:  $e$  (Sally's perceptual experience) and  $B$  (Sally's belief that there is a red patch). In this case,  $e$  is the sole member of Sally's reasons,  $R$ . Given the details of this case, it is easy to see how **IB-R** can be applied to determine whether Sally's belief is based on her reasons. We can see if each  $r_i \in R$  is a direct cause of  $B$  by seeing if there are possible interventions on  $e$  that would result in changes in  $B$ . One possible intervention we could perform on  $e$  to test whether it is a direct cause of  $B$  would be to simply set the value of  $e$  to 0. Setting  $e$  to 0 in this case will lead to a change in  $B$ . Intuitively, if we remove Sally's perceptual experience (set  $e$  to 0) in this case, she will no longer believe that there is a red patch ( $B$  will change because Sally will no longer have her perceptual belief). So, the first condition of **IB-R** is met. Next, we need to determine if the second condition of **IB-R** is satisfied. In light of the fact that changing the actual value of  $e$  results in the actual value of  $B$  changing, it is clear that each  $r_i \in R$  is an actual cause of  $B$  in this case. Of course, since  $e$  is the only direct cause of  $B$  in this case, the third condition of **IB-R** is satisfied as well. Consequently, **IB-R** is satisfied in this case. Thus, **IB-R** yields the intuitively correct result that Sally's belief that there is a red patch is based on her reasons.

Another simple case that is only slightly more complex than the previous example is a case involving simple inferential reasoning. Consider the following case (Case 2):

Sally justifiedly believes that  $p$  and she justifiedly believes that  $(p \rightarrow q)$ . From these two beliefs, she infers  $q$ .

There are two ways that we can understand this case depending on whether we think that in order for  $p$  and  $(p \rightarrow q)$  to propositionally justify  $q$  for Sally she has to recognize *modus ponens* as a valid inference form or if her justifiedly believing  $p$  and  $(p \rightarrow q)$  is sufficient for  $q$  to be propositionally justified for Sally.<sup>21</sup> For our examination of this case, we will assume that Sally needs to recognize *modus ponens* as a valid inference form in order to have propositional justification for  $q$  and that she has this recognition.<sup>22</sup> Additionally, we will assume that there are no causally relevant factors for Sally's belief that  $q$  aside from Sally's belief that  $p$ , her belief that  $(p \rightarrow q)$ , and her recognition of *modus ponens* as a valid inference form. So, in this case the variables in the variable set  $\mathbf{V}_2$  will be  $p$  (Sally's belief that  $p$ ),

<sup>21</sup> One who thinks that recognition of *modus ponens* as a valid inference form is necessary for  $q$  to be propositionally justified for Sally need not think that Sally has to have the ability to articulate this or even that she has the concepts of 'validity' or '*modus ponens*'. Instead, one might think that the sort of recognition needed is simply the awareness that *modus ponens* is a good way to reason, that is to say, an awareness that when one is presented with  $p$  and  $(p \rightarrow q)$  it is reasonable to infer  $q$ .

<sup>22</sup> This case can easily be modified to accommodate the intuition that Sally does not need this sort of recognition in order to have propositional justification for  $q$  by removing Sally's recognition of *modus ponens* as a valid inference form from  $R$  and from the set of causally relevant features. In other words, the case can be modified so that the only causally relevant features are Sally's beliefs  $p$  and  $(p \rightarrow q)$ .

$c$  (Sally's belief that  $(p \rightarrow q)$ ),  $m$  (Sally's recognition of *modus ponens*), and  $B$  (Sally's belief that  $q$ ). Again, the first step in determining whether  $B$  is based on  $R$  is to determine if each of the members of  $R$  ( $p$ ,  $c$ , and  $m$ ) is a direct cause of  $B$ . Presumably, in this case each member of  $R$  will be a direct cause of  $B$  because there are interventions that we could perform on  $p$ ,  $c$ , or  $m$  that would result in a change in  $B$ . For instance, if we hold fixed  $c$  and  $m$  while lowering the value of  $p$  to 0 via an intervention, Sally would no longer have any reason to infer  $q$  (so the value of  $B$  would change) because in the case described she is making an inference from  $p$  and  $(p \rightarrow q)$  to  $q$  via *modus ponens*. The same would apply *mutatis mutandis* for  $c$  as well as  $m$ . So, in this case each member of  $R$  is a direct cause of  $B$ . Moreover, each member of  $R$  is an actual cause of  $B$ . That is, there are interventions on the actual values of the members of  $R$ , each of which results in the actual value of  $B$  changing. One such intervention involves holding the values of the other variables fixed at a value within their redundancy ranges while lowering the actual value of  $p$  to 0. With the value of  $p$  set at 0 the value of  $B$  will change to 0. Intuitively, if Sally does not believe that  $p$  she will no longer infer  $q$  in this case. The same considerations apply *mutatis mutandis* for both  $c$  and  $m$ . In other words, it is also the case that if Sally does not believe  $(p \rightarrow q)$ , she will not infer  $q$ ; and if Sally does not recognize *modus ponens* as a valid inference form, she will not infer  $q$ . Since the only direct causes of Sally's believing that  $p$  are members of  $R$ , it is obvious that the third condition of **IB-R** is satisfied in this case. Thus, again **IB-R** yields the intuitive result: Sally's belief that  $q$  is based on her reasons.

Before moving on to examining different kinds of cases, it will be illuminating to consider how **IB-R** applies to more complicated instances of inferential reasoning. Specifically, we will consider what we should say about a case where someone uses a bad inference rule in coming to a belief and what we should say about the basis for a belief that is the end result of a chain of reasoning, if we accept **IB-R**. We will begin by considering the former kind of case (Case 3):

Sally justifiably believes that  $p$  and she justifiably believes that  $(p \rightarrow q)$ . From these two beliefs, she infers that  $q$ . However, Sally is not relying on *modus ponens* to form her belief. Instead she is relying on some other (invalid) inference rule,  $X$ .<sup>23</sup>

Just like the previous case there are two ways to understand this sort of situation. One way is to think that in order for Sally's justifiably believing  $p$  and  $(p \rightarrow q)$  to propositionally justify  $q$  for her she has to recognize *modus ponens* as a valid inference form. Another way is to think that the fact that Sally justifiably believes  $p$  and she justifiably believes  $(p \rightarrow q)$  is itself sufficient for  $q$  to be propositionally justified for her. If we opt for the latter understanding, then there is no relevant difference between this case and cases where someone reasons via *modus ponens*. So, given this understanding of the situation, **IB-R** yields the result that Sally's belief is based on her reasons, which is intuitively correct.

<sup>23</sup> The exact nature of inference rule  $X$  is not important for our purposes. All that matters for our concerns is that  $X$  is invalid, but it does lead Sally to infer  $q$  from  $p$  and  $(p \rightarrow q)$  in this particular case.

Now, if we assume that Sally's recognition of *modus ponens* is part of her propositional justification for  $q$ , then Sally's belief that  $q$  is not based on her reasons. So, given this assumption about the reasons that propositionally justify Sally's belief a satisfactory account of the basing relation should rule out the possibility of Sally's belief that  $q$  being based on her reasons in the sense required for doxastic justification. This is exactly what **IB-R** does. Let us examine how it does so. In this case the variables in the relevant variable set  $V_3$  are:  $R$  (Sally's reasons),  $X$  (Sally's applying inference rule  $X$ ), and  $B$  (Sally's belief that  $q$ ). In this case, the members of  $R$  are:  $p$  (Sally's belief that  $p$ ),  $c$  (Sally's belief that  $(p \rightarrow q)$ ), and  $m$  (Sally's recognition of *modus ponens*). Given that Sally does not rely on recognition of *modus ponens* when making her inference, if we hold the other variables in  $V_3$  fixed at values in their redundancy ranges, intervening to change the actual value of  $m$  will not result in a change in the actual value of  $B$ . That is to say, since Sally believes that  $q$  because of her beliefs  $p$  and  $(p \rightarrow q)$  and her application of inference rule  $X$ , intervening so that she recognizes that *modus ponens* is a valid inference form will not lead to a change in whether she believes that  $q$ . Therefore,  $m$  is not an actual cause of  $B$ . So, the second condition of **IB-R** is not met. Additionally, if the value of  $X$  is set to 0 while the members of  $R$  are held fixed at their actual values, Sally will no longer believe that  $p$  because she is relying on inference rule  $X$  when forming her belief. So, the third condition of **IB-R** is not satisfied in this case either. Thus, we can see that **IB-R** allows us to rule out cases of bad inference from being cases in which one's belief is based upon the reasons which propositionally justify the belief.

At this point, we will consider how **IB-R** applies to cases that involve chains of reasoning. Here is a simple case of someone progressing through a chain of reasoning (Case 4):

Sally justifiably believes that  $p$  and she justifiably believes that  $(p \rightarrow q)$ . From these two beliefs, she infers that  $q$ . Sally also justifiably believes that  $(q \rightarrow t)$ . From this belief and her new belief that  $q$ , she infers that  $t$ . Sally is utilizing the inference rule *modus ponens* for both of her inferences.<sup>24</sup>

One might worry that **IB-R** will commit us to a counter-intuitive position concerning Sally's belief that  $t$ . The reason for this worry is that the variables representing Sally's belief that  $q$  and her belief that  $(q \rightarrow t)$  will be direct causes of  $t$ , but those representing her belief that  $p$  and her belief that  $(p \rightarrow q)$  will not. So, it seems that **IB-R** saddles us with a dilemma because according to **IB-R** either Sally's belief that  $t$  is not based on her reasons because Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are not direct causes of  $t$  or Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are not part the reasons which propositionally justify Sally in believing that  $t$ . The crux of this worry is that it seems that on any viable account of justification Sally's doxastic justification for her belief that  $t$  (assuming that  $t$  is propositionally justified) will be dependent upon her justifiably believing that  $p$  and her justifiably believing that

<sup>24</sup> It is stipulated in this case that Sally is using *modus ponens* in order to make it clear that this is not a case of bad inference like the preceding example.

$(p \rightarrow q)$ , however, **IB-R** does not seem to allow for the proper justificatory connection between these beliefs.

Although at first glance it may seem that **IB-R** commits us to a counter-intuitive position concerning chains of reasoning, the situation is not as bad as it seems. It is true that in the above case **IB-R** commits us to claiming that neither Sally's belief that  $p$  nor her belief that  $(p \rightarrow q)$  are direct causes of Sally's belief that  $t$ . It is also true that if we accept both **IB-R** and that Sally's belief that  $t$  is based on her reasons, we have to say that Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are not part of Sally's reasons, at least insofar as we are referring to the variables that are members of  $R$  as her 'reasons'. However, the crucial point to recognize is that **IB-R** only commits us to claiming that Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are not part of  $R$ , it does not commit us to claiming that they are not part of Sally's reasons simpliciter. We can claim that Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are part of her reasons for believing that  $t$  in a sense that is typical of foundationalist theories of justification. Sally's beliefs that  $p$  and that  $(p \rightarrow q)$  are her reasons for believing that  $q$ , which is itself one of Sally's reasons for believing that  $t$ . So, **IB-R** allows us to understand the dependence of Sally's belief that  $t$  on her beliefs that  $p$  and that  $(p \rightarrow q)$  in the following intuitively correct fashion: Sally's belief that  $t$  can be doxastically justified only if it is based on her reasons for  $t$ , which are her beliefs that  $q$  and that  $(q \rightarrow t)$ , and those beliefs are themselves doxastically justified. Sally's belief that  $q$  can be doxastically justified only if it is based on her reasons for  $q$ , which are her beliefs that  $p$  and that  $(p \rightarrow q)$ . So, although **IB-R** does not allow us to claim that Sally's belief that  $t$  is based on her beliefs that  $p$  and that  $(p \rightarrow q)$ , it is clear that **IB-R** does not conflict with the intuitively correct assumption that Sally's justification for her belief that  $t$  is dependent upon her beliefs that  $p$  and that  $(p \rightarrow q)$ . Thus, chains of reasoning do not pose a problem for **IB-R**.

Another important kind of case that should be considered is one in which someone possesses several independent reasons for a belief which together propositionally justify the belief. The following is such a case (Case 5):

Sally has independent reasons  $r_1$ ,  $r_2$ , and  $r_3$  for believing that  $p$ . These reasons are such that together they propositionally justify  $p$ , but separately or even in pairs they do not propositionally justify  $p$ . Additionally, Sally only believes that  $p$  because she has all three of  $r_1$ ,  $r_2$ , and  $r_3$ .

In this case the variables in the relevant variable set  $V_5$  are:  $R$  (Sally's reasons) and  $B$  (Sally's belief that  $p$ ). In this case the members of  $R$  are:  $r_1$ ,  $r_2$ , and  $r_3$ . Each member of  $R$  is a direct cause of  $B$ . This is clear because Sally only believes that  $p$  because she has all three reasons. So, in this case if the value of  $r_1$  were changed while the values of  $r_2$  and  $r_3$  were held fixed, the value of  $B$  would change. The same would apply *mutatis mutandis* for  $r_2$  as well as  $r_3$ . Moreover, each member of  $R$  is an actual cause of  $B$ . That is, there are interventions on the actual values of each of the members of  $R$ , each of which results in the actual value of  $B$  changing. Holding the values of the other variables ( $r_2$  and  $r_3$ ) fixed at values within their redundancy ranges there is an intervention that can be performed on  $r_1$ , which changes its actual value, resulting in a change in the actual value of  $B$ . One such intervention is to lower the actual value of  $r_1$  to 0. With the value of  $r_1$  set to 0 the value of  $B$  will

change to 0. Simply put, the idea is that if Sally does not have  $r_1$  as a reason in this case she will no longer believe that  $p$ . The same considerations apply *mutatis mutandis* for both  $r_2$  and  $r_3$ . Further, the third condition of **IB-R** is satisfied in this case because there are no direct causes of  $B$  other than the members of  $R$ . So, **IB-R** yields the result that Sally's belief that  $p$  is based on her reasons, which is intuitively correct.

Now we will examine a slightly more complex case, which involves causally relevant factors in addition to  $S$ 's reasons (Case 6):

Sally has reasons  $r_1 \dots r_n$  for believing that  $p$ . However, Sally also really wants  $p$  to be true. Her desire for the truth of  $p$  is such that she cannot help but engage in wishful thinking when considering the truth of  $p$ .<sup>25</sup>

Unfortunately, this case is probably similar to our situation with regard to several of our beliefs—we have reasons in support of the belief, but we also have mechanisms that are not reasons in support of our belief that are causally relevant to our having the belief. Thus, it will be especially important to get clear on how **IB-R** handles such cases.

Given the features of this case the variables in the relevant variable set  $V_6$  are:  $r_1 \dots r_n$  (for each of Sally's reasons),  $w$  (Sally's wishful thinking mechanism), and  $B$  (Sally's belief that  $p$ ). Let us assume for simplicity that in this case Sally's reasons,  $R$ , consist only of  $r_1$ . Further, let us assume that  $r_1$  and  $w$  are both direct causes of  $B$  and they are both actual causes of  $B$  as well. Finally, let us assume that both  $r_1$  and  $w$  are sufficient on their own for  $S$  to believe that  $p$ . That is to say, if Sally did not have her reasons for  $p$ , she would still believe  $p$  because of her wishful thinking; and if Sally did not engage in wishful thinking, she would still believe that  $p$  because of her reasons. To be clear, we are construing the above example as one in which both Sally's reasons and her wishful thinking are direct causes and actual causes of her believing that  $p$  and both her reasons and her wishful thinking are individually sufficient to cause Sally to believe that  $p$ . This way of construing the example will help to demonstrate an advantage **IB-R** has over some other causal accounts of the basing relation. Specifically, this case creates a problem for causal accounts of the basing relation that rely upon simple counterfactual accounts of causation. According to simple counterfactual accounts of causation, in this case Sally's reasons fail to be a cause of her believing that  $p$  because it is not the case that if Sally did not have her reasons, she would not believe that  $p$ . The problem for these accounts stems from the fact that Sally's belief that  $p$  is over-determined by her reasons and her wishful thinking. Since Sally's belief is over-determined, these accounts yield the counterintuitive result that her belief that  $p$  is not based on her reasons because Sally's reasons are not a cause of her believing that  $p$  according to these accounts. **IB-R** does not share this problem. All three conditions of **IB-R** are satisfied in this case because we have stipulated that  $r_1$  and  $w$  are direct causes of  $B$  and that  $r_1$  and  $w$  are actual causes of  $B$ . Additionally, we have stipulated that

<sup>25</sup> In this case we will make use of wishful thinking, but the wishful thinking component could be replaced with any feature(s) of one's situation that is not a reason for/against the truth of her belief that may, nonetheless, be causally relevant to her holding her belief.

Sally's reasons are sufficient for her believing that *p*. In other words, it is not the case that Sally's failing to engage in wishful thinking (setting *w* to 0), while continuing to have her reasons (holding *r<sub>i</sub>* fixed at its actual value) will result in her not believing that *p*. **IB-R** yields the intuitively correct result that Sally's belief that *p* is based on her reasons even though her belief is over-determined by her reasons and her wishful thinking, but causal accounts of the basing relation which utilize a simple counterfactual account of causation do not yield this result. Therefore, in this case an advantage of **IB-R** over causal accounts of the basing relation that utilize simple counterfactual accounts of causation is made clear.<sup>26</sup>

As can be seen from our analysis of this case, **IB-R** yields the result that whether or not someone's beliefs are based on her reasons depends on the strength of the causal relationship between her reasons and her beliefs. Thus, just because an epistemically suspect factor plays a causal role in someone's believing as she does, it does not necessarily mean that her belief is not based on her reasons. Not only is this result intuitively correct, if this result were not correct it would spell trouble for most, if not all, ordinary believers because it is likely that there are epistemically suspect factors causally influencing a large number (perhaps the vast majority) of our beliefs to at least some degree.

Now that we have considered **IB-R** and seen how it applies to a number of cases, we will turn our attention to examining objections that have been raised against causal accounts of the basing relation and to evaluating their effectiveness as objections to **IB-R**.

## 4 Objections

### 4.1 Causal deviancy

One problem facing causal analyses in general is how to deal with deviant causal chains. In order for a causal analysis of any phenomenon to be successful it must be able to differentiate between instances of deviant and non-deviant causal chains in a principled manner. Thus, causal accounts of the basing relation face the challenge of providing a way of ruling out cases of deviant causation from counting as instances where one's beliefs are based on one's reasons while not ruling out cases of non-deviant causation in the process. Many causal accounts of the basing relation fail to meet this challenge.<sup>27</sup> Here another advantage of **IB-R** over many other causal accounts of the basing relation becomes clear. Ruling out causally deviant chains

<sup>26</sup> Other cases that one might worry about are cases where one cause of *S*'s belief preempts another cause and cases where one cause of *S*'s belief trumps another cause. The interventionist account of causation provides **IB-R** with the conceptual tools necessary for yielding the intuitive results in these sorts of cases as well. The interested reader is advised to consult Woodward (2003) for an explanation of how the interventionist account of causation handles these sorts of cases, particularly, pp. 77–82.

<sup>27</sup> Some, such as Moser (1989), simply build in a 'nondeviant' clause in their accounts. However, simply stipulating that the causal process resulting in *S*'s belief that *p* must not be deviant in order for the belief to be based on *S*'s reasons merely acknowledges the problem posed by causal deviancy without making any progress toward solving it. Others such as Pollock and Cruz (1999) merely note that causal deviancy poses a problem.

from counting as instances where a belief is based on reasons is something that **IB-R** does extremely well. In order to illustrate this feature of **IB-R** we will consider three cases, which involve distinct forms of causal deviancy, and how **IB-R** delivers the intuitively correct results in each case.

We will begin with an example of causal deviancy from Alvin Plantinga:

Suddenly seeing Sylvia, I form the belief that I see her; as a result, I become rattled and drop my cup of tea, scalding my leg. I then form the belief that my leg hurts; but though the former belief is a (part) cause of the latter, it is not the case that I accept the latter on the evidential basis of the former.<sup>28</sup>

Let us consider how **IB-R** applies to Plantinga's example. The intuitive thought that causal accounts of the basing relation must accommodate is that although Plantinga's belief that he sees Sylvia is a cause of his belief that his leg hurts (it is part of the causal chain leading to that belief), his belief that his leg hurts is not based on his belief that he sees Sylvia. According to **IB-R**, in order for Plantinga's belief that his leg hurts to be based on his belief that he sees Sylvia the latter belief must be a direct cause of the former. However, in Plantinga's example his belief that he sees Sylvia is not a direct cause of his belief that his leg hurts. Holding the value of all other variables, particularly the scalding of his leg, fixed in this situation and performing interventions on Plantinga's belief that he sees Sylvia will not result in his confidence for the belief that his leg hurts changing. So, Plantinga's belief that his leg hurts is not based on his belief that he sees Sylvia.

On the other hand, given the details of this example, it is plausible that **IB-R** will commit us to claiming that Plantinga's belief that his leg hurts is based on his reasons in the sense required for doxastic justification. Since Plantinga's reasons for his belief that his leg hurts (*B*) consists of his experience of his leg being scalded (*s*), the first two conditions of **IB-R** are satisfied. That is, and *s* is a direct cause and an actual cause *B*. Further, there are no direct causes of Plantinga's belief other than his experience of his leg being scalded, so the third condition of **IB-R** is clearly satisfied in this case. Thus, according to **IB-R**, Plantinga's belief that his leg hurts is based on his reasons. This is the case regardless of the fact that Plantinga's belief that he sees Sylvia is part of the causal chain leading to his experience of his leg being scalded and to his belief that his leg hurts. Intuitively, in Plantinga's example his belief that his leg hurts is not based on his belief that he sees Sylvia, it is based on his experience of his leg being scalded; and this is exactly what **IB-R** says.

Another example of causal deviancy that we should consider is the kind that occurs in the following sort of case (Case 7):

Sally has reasons  $r_1 \dots r_n$  for believing that *p*. Each of these reasons is a direct cause of Sally's belief that *p* and each is an actual cause of Sally's believing that *p*. However, the neurological state(s) that realizes Sally's having  $r_1 \dots r_n$  causes neurological state *X*, which does not realize a mental state. *X* causes the neurological state that realizes Sally's believing that *p* and the neurological

<sup>28</sup> Plantinga (1993, p. 69).



pathway from the neurological realizer(s) of  $r_1 \dots r_n$  to X to the realizer of Sally's believing that p is deviant.

In this case **IB-R** gives us the result that Sally's belief is based on her reasons. The three conditions of **IB-R** are satisfied in this case, the first two conditions by stipulation and the third by default since there are no direct causes of Sally's belief besides her reasons. So, even though the neurological process that leads from Sally's reasons to her believing that p is causally deviant, **IB-R** commits us to claiming that Sally's belief that p is based on her reasons. This result is intuitively correct. After all, the mental states that cause Sally's belief that p are mental states that constitute her reasons. In general, it does not seem to matter that the realizers of one's mental states behave in strange ways. All that matters for basing one's beliefs on one's reasons is that one's mental states, those which are one's reasons, bear the appropriate causal relationship to that belief. Thus, this form of causal deviancy does not pose a problem for **IB-R**.

A final example of causal deviancy, which is often considered particularly problematic for causal accounts of the basing relation, is exemplified in the following sort of case (Case 8):

Sally has reasons  $r_1 \dots r_n$  for believing that p. These reasons cause Sally to think about how much she likes butterflies (which has no justificatory relationship to p nor to any of  $r_1 \dots r_n$ ). Thinking about how much she likes butterflies causes Sally to believe that p.

It is clear in this case that because of the deviancy at work Sally's belief that p is not based on her reasons in the sense required for doxastic justification. That is to say, Sally's belief that p is not based on the reasons which propositionally justify her belief that p. It is easy to see how **IB-R** yields this intuitively correct result. The first condition of **IB-R** requires that each of Sally's reasons for believing that p must be direct causes of p. In this case, none of Sally's reasons is a direct cause of her belief that p. Instead, Sally's reasons causally contribute to her believing that p by directly causing her mental state of thinking about how much she likes butterflies, which directly causes her belief that p. So, the first condition of **IB-R** is not satisfied in this case. Moreover, the third condition of **IB-R** is also not satisfied. If Sally's reasons are held fixed and her butterfly belief is removed (we set the value of this direct cause to 0), Sally will not believe that p. Thus, in this case Sally's belief is not based on her reasons in the sense required for doxastic justification. Again, **IB-R** provides the intuitively correct response to this instance of causal deviancy.

Although there are a variety of ways that causal deviancy might affect the basing of one's beliefs, consideration of the three examples examined above suggest that many of the ways that causal deviancy affects basing seem to share a common feature. When causal deviancy leads us to conclude that a belief of S's is not based on her reasons, it is often because we recognize that her reasons are not the key causal feature of S's having that belief, instead the key causal feature is whatever makes the causal chain go wrong. That is to say, in these sorts of cases we think that S's belief is not based on her reasons because her belief is directly caused by something other than her reasons. Given the plausible assumption that direct

causation is the central feature that guides our judgments in these cases, it is reasonable to think that **IB-R** will provide the correct response to other cases of causal deviancy.

#### 4.2 Gypsy lawyers and ill-motivated politicians

A second objection that has been pressed against causal accounts of the basing relation has been advanced by Keith Lehrer and others by proposing the case of the Gypsy Lawyer and other similar purported counterexamples.<sup>29</sup> Lehrer's example is a situation in which a gypsy lawyer believes that his client is innocent of a particular crime because the tarot cards say that the client is innocent. After the lawyer's consultation of the tarot cards he comes to possess other evidence for thinking that his client is innocent and he is aware that this evidence is good evidence for believing that his client is innocent. However, it is the lawyer's trust in the tarot cards that caused and is sustaining his belief that his client is innocent, not his evidence. Lehrer explains the general nature of his example in this way:

In my example a man comes to believe something and continues to believe it because of groundless superstition...he uncovers reasons for the belief that give him knowledge. But, these reasons do not potentially explain his belief, because he would not hold the belief for those reasons if he were to become doubtful of his superstitious reasons for belief.<sup>30</sup>

Lehrer's example is a purported counterexample to causal accounts of the basing relation if we assume that a necessary condition for knowing that *p* is that one's belief that *p* be based on the reasons which propositionally justify the belief. In Lehrer's example the gypsy lawyer is supposed to come to know that his client is innocent, so his belief that his client is innocent must be based on his reasons. However, in Lehrer's example the reasons that make believing that the client is innocent propositionally justified for the lawyer are not causally relevant to his having that belief.

Lehrer's example has failed to convince supporters of causal accounts of the basing relation that there is a genuine problem.<sup>31</sup> The key problem with Lehrer's example is that it is not clear why we should think that the gypsy lawyer *knows* that his client is innocent.<sup>32</sup> The gypsy lawyer has good reasons for thinking that his client is innocent, however, it is far from clear that merely having good reasons for believing *p* is enough for one's belief that *p* to be justified in the way required for knowledge. It is perfectly consistent to think that the gypsy lawyer, while being in possession of good reasons, fails to adequately use those reasons; and so, fails to

<sup>29</sup> See Lehrer (1971) for the original presentation of this purported counterexample. Also, see Lehrer (1974) and Korcz (2000) for variations on this example.

<sup>30</sup> Lehrer (1971, p. 311).

<sup>31</sup> For example see Goldman (1979) and Pollock and Cruz (1999) who simply deny that the example shows what Lehrer claims it does. Also, see Audi (1983) who argues more extensively against the effectiveness of Lehrer's example.

<sup>32</sup> Audi (1983), Goldman (1979), and Pollock and Cruz (1999) all maintain that the gypsy lawyer fails to know that his client is innocent.

know that his client is innocent. After all, the intuition that motivates the necessity of basing for justified beliefs is that cases like this are not instances of knowledge. At the very least it is safe to conclude that Lehrer's example is not an obvious counterexample to causal accounts of the basing relation. So, instead of defending **IB-R** from the purported conclusions of this problematic example, let us examine an example that Jonathan Kvanvig maintains supplies much needed support to Lehrer's example.<sup>33</sup>

Although Kvanvig admits that Lehrer's example has failed to persuade, he claims to provide a strong argument that effectively demonstrates the truth that Lehrer attempts to illustrate with the Gypsy Lawyer example. Kvanvig's argument relies on the assumption that "a proper account of justified belief will be extendable to an account of justified behavior."<sup>34</sup> The idea is that the distinction between having good reasons to believe and believing on the basis of those reasons is analogous to the distinction between having good reasons to act and performing an action on the basis of those reasons. Kvanvig argues that causal accounts of the basing relation are inadequate because they cannot make sense of this feature of the justification of behavior. Kvanvig's argument utilizes an example of someone who is running for Congress for irrational reasons (Ill-motivated Politician). Although the politician, Jim, is actually motivated by irrational reasons, he offers rational reasons for his behavior when asked why he is running for Congress. As time goes by Jim becomes aware of this feature of himself. Additionally, Jim realizes that the reasons which he has been giving others to justify his running for Congress are in fact good reasons for him to do so. Kvanvig describes Jim's situation after his newfound self-awareness in the following manner:

Jim has come to realize his true motives. He has come to realize that the reasons he has given for running are not what brings him to run for Congress...Jim comes to realize that the reasons he has been offering for his behavior (i) did not originally prompt the behavior, (ii) have not, in the past, sustained the behavior, and (iii) do not now sustain the behavior. Regarding this third fact, what Jim realizes is that he is so constituted at present that the reasons he has offered do not even enhance the probability of his running, even if we were to control for the causal force of his irrational desire. Upon confronting these rather disturbing facts, Jim then reasons as follows: "the inadequate motivations of both past and present are regrettable and everything possible ought to be done to alter them; but, until this alteration can be accomplished, everything possible ought to be done to maintain some motivation or other to keep running for Congress since, after all, it is nonetheless true that I am extraordinarily good at convincing others of correct policy, that I am best qualified to serve the constituents of this district, and if persons were to attempt to quit doing everything which is done for inadequate reasons, not (as) much good would be done." So, Jim concludes, he ought to

---

<sup>33</sup> Kvanvig (2003)

<sup>34</sup> Kvanvig (2003, p. 50).

do all in his power to keep the race for Congress alive in spite of his bad motivations.<sup>35</sup>

Kvanvig divides the events of his example into three stages: stage one—Jim’s situation prior to discovering that his running for Congress is not justified, stage two—Jim’s situation after he realizes his real motivations for running for Congress, but prior to his formulating *R* (good reasons for his running for Congress), and stage three—Jim’s situation after formulating *R*. Kvanvig claims that as Jim progresses through these stages he makes rational progress, which is something that must be explained. Kvanvig explains this fact by insisting that Jim transitions in this case from merely “performing a justified action to justifiably performing that action.”<sup>36</sup> According to Kvanvig the most natural assessment of Ill-motivated Politician is to claim that Jim lacks “the kind of rationality which implies that he is justifiably running for Congress” before he constructs *R*, however, once Jim constructs *R*, “his running is rational in a way which implies that it is perfectly justified.”<sup>37</sup> Kvanvig points out that this way of assessing Ill-motivated Politician is not available to those who accept a causal account of the basing relation because they will be committed to claiming that Jim is not justifiably running for Congress at any of the three stages of this case. Kvanvig concludes that since there is no satisfactory account of Jim’s rational progress in Ill-motivated Politician available to those who accept a causal account of the basing relation, causal accounts of the basing relation are not satisfactory.

Admittedly, Kvanvig’s objection to causal accounts of the basing relation cannot be simply dismissed as unconvincing in the way that many think Lehrer’s example can. However, there are reasonable responses available to those who accept **IB-R** (or any other causal account of the basing relation for that matter). The first way we might respond to Kvanvig’s argument is to note that it is not clear that we should accept his assumption that an adequate account of justified belief should extend to justified behavior. There are a variety of factors that are relevant to the justification of one’s actions such as one’s desires, beliefs, access to resources, commitments to others, goals, etc. It is not clear that these same factors are also relevant to the justification of one’s beliefs. For example, in a Buridan’s Ass situation where there are two actions one may perform and no reason to prefer one action over another, but good reasons to prefer some action to no action, it seems that *S* will justifiably perform whichever of the two actions she does in fact perform. However, there does not seem to be an analogous situation in regard to justified belief. Just because *S* is in a situation where she has no reason to think that *p* is more likely to be true than  $\sim p$  and she has good reasons to prefer adopting some attitude toward *p* other than withholding; it clearly does not follow that *S* will justifiably believe whichever of *p* or  $\sim p$  that she actually ends up believing. Although these considerations do not demonstratively show that a satisfactory account of justified belief does not need to be extendable to an account of justified behavior, they do raise at least some

<sup>35</sup> Kvanvig (2003, pp. 50–51).

<sup>36</sup> Kvanvig (2003, p. 61).

<sup>37</sup> Kvanvig (2003, p. 61).

doubt concerning Kvanvig's assumption that a satisfactory account of justified belief must be extendable in the way he claims.

A second way that we may respond to Kvanvig's objection is to argue that **IB-R**, or causal accounts of the basing relation in general, does not conflict with a plausible account of how Jim has made rational progress in Ill-motivated Politician. It is plausible that what accounts for Jim's rational progress in Kvanvig's example is that after formulating *R* Jim has reasons that could make his running for Congress justified. This is not to agree with Kvanvig's claim that Jim's running for Congress is justified after he formulates *R*. Instead, once Jim formulates *R* he is in possession of a necessary component for his running for Congress to be justified. Jim still needs to base his running for Congress on *R* for his running to be justified, but that possibility is not accessible to Jim until he reaches stage three of Kvanvig's case.<sup>38</sup> So, one can reasonably maintain that Jim makes progress toward becoming more rational in Ill-motivated Politician without denying **IB-R**.

A third way that we may respond to Kvanvig's objection is to deny that Jim really has made rational progress regarding his actions in Ill-motivated Politician. True enough, Jim has made some gains in self-awareness in this example by realizing his true motivations for running for Congress and formulating good reasons for why he should run for Congress. However, these changes in Jim's mental states do not amount to changes in his behavior. So, much like the person who learns that helping someone would be a morally good thing to do, but helps her only because he believes it will lead to personal gain has not made progress toward behaving in a moral manner; Jim recognizes reasons that would make his action justified, but he does not make progress toward justifiably acting because he continues to act for the wrong reasons. As a result, it is plausible to maintain that Jim does not make rational progress of the sort that is supposed to be problematic for **IB-R** or causal accounts of the basing relation in general, to explain.

So, in Ill-motivated Politician either Jim's actions after he reaches stage three are based on *R* or they are not. If they are, then his actions are rational; if not, then his actions are irrational. This is the intuitively correct assessment of Jim's situation. While it remains an open question whether or not Jim has made some other sort of rational progress throughout his process of self-realization, this sort of rational progress, if it occurs, is not something that a satisfactory account of the basing relation needs to explain.

### 4.3 Mutually supporting beliefs

Another objection that can be raised against causal accounts of the basing relation stems from the fact that there can be mutually supporting beliefs. Cases like the following are not uncommon (Case 9):

Sally is working on a crossword puzzle. She believes that one-across is "acidic" in part because she believes that one-down is "airplane". Additionally, she

<sup>38</sup> Kvanvig considers and rejects a number of attempts to explain Jim's rational progress. However, he does not consider this explanation of Jim's rational progress. Furthermore, his arguments against the attempts that he rejects do not impugn this response.

believes that one-down is “airplane” in part because she believes that one-across is “acidic”.<sup>39</sup>

One might take examples of this sort to pose a problem for causal accounts of the basing relation because intuitively Sally’s beliefs that one-across is “acidic” and that one-down is “airplane” can provide epistemic support for one another and it seems plausible that the two beliefs can be at least partly based on one another. The problem that this seems to pose for causal accounts of the basing relation is that it appears that given a causal account of the basing relation the only way that two beliefs can be partly based on one another is if causal asymmetry is denied. Thus, one might think that supporters of causal accounts of the basing relation face a dilemma: deny that there can be mutually supporting beliefs or deny causal asymmetry.

Fortunately, there are at least two plausible ways of responding to this purported dilemma. The first is simply to embrace the second horn by denying causal asymmetry.<sup>40</sup> By denying causal asymmetry causal accounts of the basing relation can account for Sally’s two beliefs being partly based on one another by allowing that the two beliefs are partial causes of one another. Although denying causal asymmetry has its costs, doing so is not clearly a mistake.<sup>41</sup> There are cases which at least seem to be instances of symmetric causation such as when two people are sitting on a see-saw and their sitting keeps the see-saw balanced or when two playing cards keep one another standing by leaning on one another as in a house of cards trick.<sup>42</sup> These sorts of cases have led some, such as John Pollock (1976), to argue that any acceptable account of causation ought to at least allow for instances of symmetric causation. Others have argued that at the very least symmetric causation cannot be ruled out a priori.<sup>43</sup> So, embracing the second horn of this purported dilemma does not seem to be an obvious mistake.

Now, if one is loathe to deny causal asymmetry, there is another way to endorse a causal account of the basing relation and respond to this purported dilemma. This second way of responding involves arguing that the purported dilemma is not a genuine dilemma. That is to say, one might plausibly respond to this purported dilemma by arguing that causal accounts of the basing relation can allow that Sally’s two beliefs are partly based on one another without denying causal asymmetry. One way to go about this is to argue that what one says about other cases of seemingly symmetric causation applies to Sally’s beliefs. Recall the house of cards example mentioned above. When one is constructing a house of cards there will be instances where two playing cards remain standing upright because they lean on one another. The two playing cards are clearly exerting some causal influence on one another. Those who maintain that causation is asymmetric must account for the

<sup>39</sup> Thanks to an anonymous reviewer for mentioning this sort of example and for pressing the objection that follows.

<sup>40</sup> It should be noted that the interventionist account of causation leaves open the question of whether causation is asymmetric.

<sup>41</sup> See Frankel (1986) for an explanation of some of the costs thought to be associated with denying causal asymmetry.

<sup>42</sup> These examples are borrowed from Frankel (1986).

<sup>43</sup> See for example Hausman (1984), Lewis (1986), and Price and Weslake (2009).

causal influence that these playing cards exert on each other in some fashion.<sup>44</sup> If the two playing cards can cause one another to stand upright without violating causal asymmetry, it seems that Sally's belief that one-across is "acidic" and that one-down is "airplane" can causally influence one another without violating causal asymmetry as well. Thus, it seems that one can plausibly respond to the purported dilemma by arguing that it is not a genuine dilemma because a supporter of a causal account of the basing relation can allow for mutually supporting beliefs to be partly based on one another without denying causal asymmetry.

Since there are at least two plausible ways of responding to this purported dilemma while endorsing a causal account of the basing relation, this way of objecting to causal accounts of the basing relation is not successful.

#### 4.4 Swamp-Sally

A final objection that one might raise for causal accounts of the basing relation comes from reflection on a swampman-style case. Consider the following situation (case 10):

Sally is standing by a swamp in a thunderstorm. The swamp is struck by lightning and miraculously a complete physical duplicate of Sally appears. Swamp-Sally differs from Sally in that none of Swamp-Sally's mental states are causally connected. After her miraculous creation Swamp-Sally has enough causal contact with the world, and perhaps a linguistic community, for her mental states to acquire content. In fact her mental states acquire all of the same content as Sally's mental states at the time of the lightning strike. However, while Swamp-Sally is in contact with the world and a linguistic community in this way the causal structure of her mental states is held fixed. So, Swamp-Sally acquires no new beliefs, loses no existing beliefs, and develops no new causal connection between beliefs while her mental states are coming to have content. After Swamp-Sally's mental states have content Swamp-Sally has a conversation with Sam. During this conversation Swamp-Sally cites her belief that *r* as her reason for believing that *p*. Afterward, Sam provides Swamp-Sally with undermining evidence for *r* and Swamp-Sally abandons her belief that *p*.<sup>45</sup>

Assuming that Swamp-Sally's situation is possible, one might think that it poses a problem for causal accounts of the basing relation. The problem arises from the fact that Swamp-Sally's citing *r* as her reason for believing *p* and her abandoning her belief that *p* because of evidence which undermines *r* suggests that Swamp-Sally's belief that *p* is based on *r*, however, there is no causal connection between Sally's belief that *r* and her belief that *p*. So, one might claim that this case is a counterexample to causal accounts of the basing relation because Swamp-Sally has a belief based on reasons, but those reasons are not causes of her belief.

<sup>44</sup> See Frankel (1986) for a plausible asymmetric causal account of the influence the two cards exert on one another.

<sup>45</sup> Thanks to an anonymous reviewer for mentioning this sort of example and for pressing the objection based on the example.

At first glance Swamp-Sally seems to be a serious problem for causal accounts of the basing relation, but closer examination reveals that this is not so. To begin, it is important get clear on what sort of undermining evidence Swamp-Sally is supposed to gain in this situation. Since Swamp-Sally's belief that *p* and her belief that *r* are causally unrelated, receiving evidence that undermines *r* will not lead Swamp-Sally to give up *p*. It is reasonable to think that this undermining evidence will lead Swamp-Sally to give up *r*, but since *r* and *p* are causally unrelated, giving up *r* will not cause Swamp-Sally to give up *p*. So, the undermining evidence cannot simply be evidence that causes Swamp-Sally to give up her belief that *r*.

Another possibility is that the undermining evidence that Swamp-Sally gains from Sam is evidence that *r* does not support *p*. However, gaining evidence that *r* does not support *p* or even coming to believe that 'r does not support p' does not bear a direct causal relation to Swamp-Sally's believing or disbelieving *p*. Gaining evidence that *r* does not support *p* or even the belief that 'r does not support p' could only cause Swamp-Sally to give up her belief that *p* by interacting with something else such as her belief that 'r supports p' or her belief that 'my only reasons for *p* are *r*', etc. But, since the case is supposed to be one in which Swamp-Sally's mental states are not causally related, it seems that whatever would plausibly interact with the evidence that *r* does support *p* to cause Swamp-Sally to give up her belief that *p* cannot do so in this case because that additional mental state is causally isolated from her belief that *p*. Thus, it seems that Swamp-Sally cannot respond to the undermining evidence in a way that suggests that her belief that *p* is based on *r* unless her mental states are causally connected.

Now, one might attempt to buttress the Swamp-Sally objection by claiming that the presence of the undermining evidence brings about some sort of causal connection between Swamp-Sally's belief that *p* and her belief that *r* which was not present before she gained the evidence. The idea here would be that the presence of the undermining evidence could bring about the sort of causal connection needed for Swamp-Sally to respond to the undermining evidence in a way which suggests her belief that *p* is based on *r* without her mental states previously being causally connected. Although construing the case in this way would allow for Swamp-Sally's gaining the undermining evidence to cause her to give up her belief that *p*, it leaves us with no reason to think that Swamp-Sally's belief that *p* can be based on her belief that *r* without the two being causally connected. This is because modifying the case in this way implies that Swamp-Sally's belief that *p* and her belief that *r* must be causally connected in order for her to respond to undermining evidence for *r* in a way that suggests her belief that *p* is based on her belief that *r*. Thus, this attempt to salvage the Swamp-Sally case is unsuccessful. Therefore, it is reasonable to conclude that Swamp-Sally and similar cases do not pose a problem for causal accounts of the basing relation.

## 5 Concluding remarks

An adequate account of the basing relation is a necessary component of any adequate account of doxastic justification, which is itself a necessary component of



any adequate account of knowledge. The mechanics of the interventionist account of causation provide the tools necessary for developing an adequate account of the basing relation. **IB-R** utilizes features of the interventionist account of causation to explain the nature of the basing relation in a way that provides the intuitively correct results when applied to cases of doxastic justification including cases of overdetermination, which are problematic for some other causal accounts. Additionally, **IB-R** is able to rule out cases of causal deviancy, yet another task that other causal accounts of the basing relation have found intractable. Gypsy lawyer-style objections that are often raised against causal accounts of the basing relation are not problematic for **IB-R**, or causal accounts in general. Finally, cases of causal preemption and swampman-style cases do not pose significant difficulties for causal accounts of the basing relation like **IB-R**. In light of these facts, it is reasonable to conclude that **IB-R** deserves serious consideration as an account of the basing relation, and thus, to conclude that **IB-R** has the potential to play a significant role in any adequate account of doxastic justification, and by extension, any adequate account of knowledge.

**Acknowledgments** My thanks to Brandon Carey, Earl Conee, Andrew Moon, Bill Rowley, Philip Swenson, and an anonymous reviewer for helpful comments and discussions. I am especially grateful to Rich Feldman and Brad Weslake for comments on numerous drafts and for hours of discussion.

## References

- Audi, R. (1983). The causal structure of indirect justification. *The Journal of Philosophy*, 80, 398–415.
- Bergmann, M. (2006). *Justification without awareness*. Oxford: Oxford University Press.
- Comesana, J. (2006). A well-founded solution to the generality problem. *Philosophical Studies*, 129, 27–47.
- Comesana, J. (2010). Evidentialist reliabilism. *Nous*, 44, 571–600.
- Feldman, R., & Conee, E. (1985). Evidentialism. *Philosophical Studies*, 48, 15–34.
- Frankel, L. (1986). Mutual causation, simultaneity, and event description. *Philosophical Studies*, 49, 361–372.
- Goldman, A. (1979). What is justified belief. In G. Pappas (Ed.), *Justification and knowledge*. Dordrecht: D. Reidel.
- Hausman, D. (1984). Causal priority. *Nous*, 18, 261–279.
- Korcz, K. (1997). Recent work on the basing relation. *American Philosophical Quarterly*, 34, 171–191.
- Korcz, K. (2000). The causal-doxastic theory of the basing relation. *Canadian Journal of Philosophy*, 30, 525–550.
- Kvanvig, J. (2003). Justification and proper basing. In E. Olsson (Ed.), *The epistemology of Keith Lehrer*. Dordrecht: Kluwer.
- Lehrer, K. (1971). How reasons give us knowledge, or the case of the gypsy lawyer. *The Journal of Philosophy*, 68, 311–313.
- Lehrer, K. (1974). *Knowledge*. New York: Oxford University Press.
- Lewis, D. (1986). *Philosophical papers* (Vol. II). New York: Oxford University Press.
- Menzies, P., & Price, H. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science*, 44, 187–203.
- Moser, P. (1989). *Knowledge and evidence*. New York: Cambridge University Press.
- Plantinga, A. (1993). *Warrant: The current debate*. Oxford: Oxford University Press.
- Pollock, J. (1976). *Subjunctive reasoning*. Boston: D. Reidel.
- Pollock, J., & Cruz, J. (1999). *Contemporary theories of knowledge* (2nd ed.). Savage, MD: Rowman and Littlefield.

- Price, H., & Weslake, B. (2009). The time-asymmetry of causation. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *Oxford handbook of causation*. Oxford: Oxford University Press.
- von Wright, G. (1971). *Explanation and understanding*. Ithaca, NY: Cornell University Press.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.
- Woodward, J. (2008a). Causation and manipulability. In Zalta E. N. (Ed.), *The stanford encyclopedia of philosophy* (Winter 2008 ed.). <http://plato.stanford.edu/archives/win2008/entries/causation-mani/>.
- Woodward, J. (2008b). Cause and explanation in psychiatry: An interventionist perspective. In K. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry: Explanation, phenomenology and nosology*. Baltimore: Johns Hopkins University Press.