



## UvA-DARE (Digital Academic Repository)

### A Causal Power Semantics for Generic Sentences

van Rooij, R.; Schulz, K.

**DOI**

[10.1007/s11245-019-09663-4](https://doi.org/10.1007/s11245-019-09663-4)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Topoi

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

van Rooij, R., & Schulz, K. (2021). A Causal Power Semantics for Generic Sentences. *Topoi*, 40(1), 131–146. <https://doi.org/10.1007/s11245-019-09663-4>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



# A Causal Power Semantics for Generic Sentences

Robert van Rooij<sup>1</sup> · Katrin Schulz<sup>1</sup>

Published online: 29 July 2019  
© The Author(s) 2019

## Abstract

Many generic sentences express stable inductive generalizations. Stable inductive generalizations are typically true for a causal reason. In this paper we investigate to what extent this is also the case for the generalizations expressed by generic sentences. More in particular, we discuss the possibility that many generic sentences of the form ‘*ks* have feature *e*’ are true because (members of) kind *k* have the causal power to ‘produce’ feature *e*. We will argue that such an analysis is quite close to a probabilistic based analysis of generic sentences according to which ‘relatively many’ *ks* have feature *e*, and that, in fact, this latter type of analysis can be ‘grounded’ in terms of causal powers. We will argue, moreover, that the causal power analysis is sometimes preferred to a correlation-based analysis, because it takes the causal structure that gives rise to the probabilistic data into account.

**Keywords** Generic sentences · Causality · Semantics · Probability

## 1 Introduction

Consider the following two causal claims:

- (1) a. John’s throw of a stone caused the bottle to break.
- b. Aspirin causes headaches to diminish.

Intuitively, these statements operate on different levels: (1-a) states a causal relation between two *tokens* of events, while (1-b) states a causal relation between two *types* of events. Stating it somewhat differently, (1-a) states what is the *actual cause* of the breaking of the bottle, while (1-b) talks about causation in a *generic* fashion: it talks about tendencies. Notice that (1-b) is stated by using a *generic sentence*. In fact, it seems to express the same content as the following generic sentence:

- (2) Aspirin relieves headaches.

But if (2) expresses the same content as (1-b), this strongly suggests that also the generic sentence (2) should be given a causal analysis. The standard way to provide a causal analysis of the actual causation statement (1-a) is as something like the following counterfactual analysis (e.g., Lewis 1973a; Halpern 2016): (i) John threw the stone and the bottle broke, and (ii) had John not thrown the stone, the bottle would not have broken. Such an analysis obviously won’t do for (1-b), and neither will it do for (2). Instead, (1-b) and (2) seem to express that particular intakes of Aspirin *tend* to cause particular states of headache to go away, because of *what it is* to be Aspirin. Or, as we will say, because of the *causal power* of Aspirin to relieve headaches. This may look like a mysterious analysis, but we will show how to operationalize it such that it can be turned into a testable statement.

The proposal that we will discuss in this paper is that many more generic statements should be given a causal analysis. A causal analysis of (2) is highly natural, because ‘relieve’ is a causal verb. But many other generic statements are stated without causal verbs.

- (3) a. Tigers are striped.
- b. Birds fly.
- c. Birds lay eggs.

---

✉ Robert van Rooij  
R.A.M.vanRooij@uva.nl  
Katrin Schulz  
K.Schulz@uva.nl

<sup>1</sup> Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands

We will discuss whether they still could, or should, be given a causal analysis as well.

This paper is structured as follows: in the following section we will briefly motivate a recently proposed frequency-based descriptive analysis according to which a generic sentence of the form ‘*ks are e*’ express inductive generalizations. We don’t want to defend this analysis in that section: that would not only take too much time, but is also already done in an earlier paper (van Rooij and Schulz in press). In that section we will also discuss a conceptual problem for this frequency-based analysis: the fact that the analysis seems too extensional. In Sect. 3 we will provide a causal explanation for the descriptive analysis making use of some natural independence assumptions. We argue that the resulting causal power proposal can solve the above mentioned conceptual problem that the frequency-based analysis of Sect. 2 is too extensional. In Sect. 3 we will also argue that the proposed causal analysis of generics can be used to analyze habitual sentences and disposition ascriptions as well. In Sect. 4 we will show that once the independence assumptions of our causal derivation are given up, a causal analysis will give rise to improved empirical predictions, but the most straightforward causal analysis will also give rise to some challenges. In Sect. 5 we will argue that by a generalized causal analysis these challenges can be met. Section 6 concludes the paper.

## 2 A Probabilistic Analysis of Generics & Its Problems

Generic sentences come in very different sorts. Consider (4-a), (4-b) and (4-c).

- (4)
- a. Tigers are striped.
  - b. Mosquitoes carry the West Nile virus. (from Leslie 2008)
  - c. Wolves attack people. (from Eckhard 1999)

We take (4-a) to be true, because the vast majority of tigers have stripes. But we take (4-b) and (4-c) to be true as well, even though less than 1% of mosquitoes carry the virus and the vast majority of wolves never attack people. Most accounts of generics, if they don’t stipulate an ambiguity, start from examples like (4-a) and then try to develop a convincing story for examples like (4-b) and (4-c) from here. In van Rooij & Schulz (in press), in contrast, we took examples like (4-b) and (4-c) as points of departure and then generalized the analysis to account for more standard examples as well, in the hope that it would lead to a more uniform analysis.

What is the natural analysis of examples like (4-b)? We take this to be that:

1. it is **typical** for mosquitoes that they carry the West Nile virus, and
2. this is highly relevant information, because of the **impact** of being bitten by a mosquito when it carries the West Nile virus.

We take it that it is intuitively quite clear when one feature has a significantly higher **impact** than another. This is normally the case when the first feature gives rise to a more negative emotional reaction than the latter. We don’t have much to offer here to a quantitative measure of ‘impact’, but we think it is closely related to the notion of ‘experienced utility’ originally proposed by Bentham (1824/1987) and propagated by Kahneman and his collaborators (e.g. Kahneman et al. 1997).<sup>1</sup>

As for **typicality**, it is obviously not required for *e* to be a typical feature for *ks* that *all ks* have feature *e*. Although almost all tigers are striped, there exist albino tigers as well, which are not striped. And although ‘(be able to) fly’ is a typical feature for birds, we all know that penguins don’t have this feature. The same examples show that *e* can be typical for *k* although not only *ks* have feature *e*: cow and cats, too, can be striped, and bats fly as well. So we need a weaker notion of typicality. We take it that *distinctiveness* matters for typicality, and thus for generics. This can be illustrated by the contrast between (5-a), which is intuitively true, versus (5-b), which is false.

- (5)
- a. Lions have manes.
  - b. \*Lions are male.

One might think that (5-b) is false because only 50%, if at all, of the lions are male, which cannot be enough for a generic to be true. But that, clearly, cannot be the reason: the only lions that have manes are male lions. Thus, not even 50% of the lions have manes. Still, (5-a) is, intuitively, true. The conclusion seems obvious: (5-a) is true, because it is *distinctive* for lions to have manes, where the notion of distinctiveness shouldn’t be too strong. On a weaker analysis of ‘being distinctive’, one demands only that *in comparison with* other larger animals, many male lions have manes. Similarly, for (4-b) to be true it is at least required that compared to other insects, many mosquitoes carry the West Nile virus. To account for this comparative analysis, of distinctiveness one could make use of either a qualitative or a quantitative analysis. But because we want

<sup>1</sup> Our notion of impact is thought of as the absolute value of ‘experienced utility’. Thus, a ‘horror’-event will have a high impact. Furthermore, we think that looking at news items helps indicating what is of impact. What is typically being reported in news items are things or events that we feel have a big impact, even if they are rather uncommon.

to incorporate the importance of the second condition, *impact*, within an analysis of ‘relatively many’, it is almost mandatory to provide a *quantitative* analysis of distinctiveness.<sup>2,3</sup> If typicality reduces to distinctiveness and if we have such a quantitative analysis of distinctiveness, plus a quantitative measure of impact, we can define a measure of *Representativeness* to account for a generic sentence of the form ‘*ks* are *e*’ as  $Distinctiveness(e, k) \times Impact(e)$  (where  $Distinctiveness(e, k)$  measures the distinctiveness of *e* for *k*). Because we will argue later that typicality cannot always be reduced to distinctiveness, the representativeness of *e* for *k*,  $Repr(e, k)$ , should be defined more generally as

- $Repr(e, k) =_{df} Typicality(e, k) \times Impact(e)$ .

Then we can say that the generic sentence ‘*ks* are *e*’ is true, or acceptable, if and only if the *representativeness* of *e* for *k*,  $Repr(e, k)$ , is high.<sup>4</sup>

- ‘*ks* are *e*’ is true, or acceptable if and only if  $Repr(e, k)$  is high.

Before we concentrate on the more general notion of typicality, let us first discuss various potential measures of **distinctiveness**. To provide a quantitative analysis of what it means that feature *e* is distinctive for group *k*, i.e., that relatively many *ks* have feature *e*, there are many options open. On one natural analysis, it holds that relatively many *ks* have feature *e* if and only if the relative frequency of *ks* that are *e* is higher than the relative frequency of alternatives of *k* that are *e*. If we measure relative frequency by probability function *P*, this can be captured by the condition that  $P(e|k)$ —i.e., the conditional probability of having feature *e* given that one is a member of group or kind *k*—is higher than  $P(e|\bigcup Alt(k))$ , where  $Alt(k)$  denotes the (contextually given) alternatives to group *k*, and  $\bigcup Alt(k)$  thus denotes the set of members of any of those alternatives. For readability, we will from now on abbreviate  $\bigcup Alt(k)$  by  $-k$ . Thus, relatively many *ks* are *e* iff  $P(e|k) - P(e|-k) > 0$ . In psychology, the measure  $P(e|k) - P(e|-k)$  is called ‘contingency’ and denoted by  $\Delta P_k^e$ . This notion plays an

important role in the theory of associative learning (cf. Shanks 1995), and it is well-known that  $\Delta P_k^e > 0$  if and only if  $P(e|k) > P(e)$ , the standard notion of *relevance*.<sup>5</sup> It should be noted, however, that  $P(e|k) - P(e|-k)$  does not behave monotone increasing with respect to  $P(e|k) - P(e)$ .<sup>6</sup> So the choice between these two measures makes a difference for predictions. Notice that if we use contingency to model distinctiveness, and if also typicality reduces to it, it is predicted that the generic ‘*ks* are *e*’ is true, or acceptable, if and only if  $[P(e|k) - P(e|-k)] \times Impact(e)$  is high. This, in turn, is high iff  $P(e|k) \times Impact(e) \gg P(e|-k) \times Impact(e)$ , if ‘ $\gg$ ’ means ‘highly above’. For features with  $Impact(e) = 1$  (which we take to be the default case), these two equalities hold iff  $P(e|k) \gg P(e|-k)$  and  $P(e|k) \gg P(e)$ , respectively, meaning that a small difference between  $P(e|k)$  and  $P(e|-k)$  (or  $P(e)$ ) is not enough to make the generic true.

Other measures to account for ‘distinctiveness’ can be used as well. One natural alternative is to use the likelihood measure  $\frac{P(e|k)}{P(e|-k)}$ , or the closely related  $(\log) \frac{P(e|k)}{P(e)}$ , to provide an analysis of ‘relatively many’. Another one is  $\frac{P(e|h) - P(e|-h)}{P(e|h) + P(e|-h)}$ , which was originally proposed by Kenemy and Oppenheim (1952), and which is a strictly increasing function of the likelihood ratio. Two yet other notions that we could use are measures of *relative difference*, like  $\Delta^* P_k^e = \frac{P(e|k) - P(e|-k)}{1 - P(e|-k)}$  and  $\frac{P(e|k) - P(e)}{1 - P(e)}$ , due to Shep (1958) and Niiniluoto and Tuomela (1973), respectively. Intuitively, these latter notions measure the amount by which *k* increases the probability of *e* to the room available for increase. These notions of ‘likelihood’ and ‘relative difference’ are used frequently in diverse fields, like epidemiology, philosophy of science, cognitive science, and social psychology. In epidemiology Shep (1958) introduced his notion to measure the *susceptibility* of a population to a risk factor. In philosophy of science these measures are used to measure the *inductive support* or *confirmation* of an hypothesis by empirical evidence (e.g., Crupi et al. 2007), in social psychology they are used to measure how *stereotypical* a feature is for a group of individuals (cf. Schneider 2004), and in cognitive science they are used to measure the *representativeness*, or *typicality*, of features for concepts [e.g., Tenenbaum and Griffiths (2001); Tentori et al. (2007)]. Just as when we use ‘contingency’ to model distinctiveness, also with these other choices it is quite clear how to incorporate *Impact(e)* into the overall measure of representativeness.

<sup>2</sup> Of course, these considerations are well-known to users of decision- and game theory. who have to combine uncertainty with utility.

<sup>3</sup> This argument won’t have any force if one takes generic sentences to be *ambiguous* between majority generics like (4-a), on the one hand, and ‘striking’ generics like (4-b) and (4-c), on the other. In fact, Leslie (2008) proposed such an ‘ambiguity’-analysis. But we don’t see any empirical evidence in favor of such an ambiguity analysis, and we thus take it to be obvious that a uniform analysis is preferred. We will see that what Leslie calls majority generics fall out as a special case of our uniform analysis.

<sup>4</sup> We won’t discuss in this paper whether generics have truth-conditions, or only acceptability conditions. Another issue we won’t discuss here is whether acceptability of generics really comes with a threshold, or whether acceptability is graded, just like representativeness.

<sup>5</sup> Cohen (1999) proposed that a generic sentence of form ‘*ks* are *e*’ is true on its relative reading iff  $P(e|k) > P(e)$ , if we limit the ‘domain’ of the probability function to  $k \cup \bigcup Alt(k)$ .

<sup>6</sup> In fact,  $P(e|k) - P(e) = P(-k) \times [P(e|k) - P(e|-k)]$  (cf. Fitelson and Hitchcock 2011).

Which of all these measures is best to account for ‘distinctiveness’ in terms of which the truth, or acceptability, of a generic sentence of the form ‘*ks* are *e*’ should be evaluated? And if ‘typicality’ doesn’t always reduce to ‘distinctiveness’, how should the former notion be defined? We are not sure whether there is a once-and-for-all answer to this question and Tessler and Goodman (in press) propose (something close to) the likelihood function, while in van Rooij and Schulz (in press) we propose that typicality should be measured by a slight variant of Shep’s (1958) notion of ‘relative difference’,  $\Delta^{**}P_k^e = \frac{\alpha P(e|k) - (1-\alpha)P(e|\neg k)}{\alpha - (1-\alpha)P(e|\neg k)}$ , with  $\alpha \in [\frac{1}{2}, 1]$ . Notice that if  $\alpha = \frac{1}{2}$ ,  $\Delta^{**}P_k^e$  comes down to Shep’s notion of distinctiveness  $\Delta^*P_k^e$ , while in case  $\alpha = 1$ ,  $\Delta^{**}P_k^e$  comes down to  $P(e|k)$ .<sup>7</sup> In sum:

- $Typicality(e, k) =_{df} \frac{\alpha P(e|k) - (1-\alpha)P(e|\neg k)}{\alpha - (1-\alpha)P(e|\neg k)} = \Delta^{**}P_k^e$  with  $\alpha \in [\frac{1}{2}, 1]$ .

Two arguments were given for this choice:<sup>8</sup>

- (i) in case  $P(e|k) = 1$  and  $P(e|\neg k) \neq 1$ , the generic sentence seems to be perfect, whatever the value of  $P(e|\neg k)$  is. In contrast to the standard notion of relevance, and to that of likelihood, this comes out by using our measure of typicality for both values of  $\alpha$ .
- (ii) in case *e* is an uncommon feature, i.e. when  $P(e|\neg k)$ , or  $P(e)$ , is low, the difference between  $P(e|k)$  and  $P(e|\neg k) - P(e|k) - P(e|\neg k)$ —should be larger for the generic to be true or appropriate than when  $P(e|\neg k)$  is high, if  $\alpha = \frac{1}{2}$ .<sup>9</sup>

<sup>7</sup> In this formulation,  $\alpha$  is just an extra contextually given free parameter. Arguably, however, one can derive the value of  $\alpha$ , by assuming that  $\alpha = \frac{P(k)}{P(k)+P(\neg k)}$ . It follows now that in case  $P(\neg k) = 0$ —i.e. when  $\bigcup Alt(k) = \emptyset$ —ends up being 1 and  $\Delta^{**}P_k^e$  comes down to  $P(e|k)$ . If we assume additionally that the tokens of the alternative kinds are chosen such that  $P(\bigcup Alt(k)) = P(\neg k) = P(k)$ , in case  $Alt(k) \neq \emptyset$ , it will also hold that  $\alpha \in \{\frac{1}{2}, 1\}$ .

<sup>8</sup> There is an argument for assuming that  $\alpha = \frac{P(k)}{P(k)+P(\neg k)}$  as well, though. Suppose that the vast majority of members of  $\bigcup Alt(k)$  are of kind *k'* and that  $P(e|k')$  is slightly higher than  $P(e|k)$ . If we don’t control for the number of tokens of alternative kinds, or types, we take into account, ‘*ks* are *e*’ will be predicted to be false, even if for most  $k'' \in Alt(k)$   $P(e|k) \gg P(e|k'')$ . But that seems wrong. One way to predict correctly would be to count not *all tokens* of the alternatives types, but rather equally many tokens of each alternative type such that we look at as many tokens if we look at the tokens of all these types together as there are tokens of *k*. Thus, it is important that we control for the number of tokens of alternative kinds, and the demand that  $P(\bigcup Alt(k)) = P(\neg k) = P(k)$  is a special case of this.

<sup>9</sup> For instance, in case  $P(e|\neg k) = 0.9$ , the value of  $\frac{P(e|k) - P(e|\neg k)}{1 - P(e|\neg k)}$  is  $10 \times [P(e|k) - P(e|\neg k)]$ , while if  $P(e|\neg k) \approx 0$ , the value of  $\frac{P(e|k) - P(e|\neg k)}{1 - P(e|\neg k)}$  is just  $P(e|k) - P(e|\neg k)$ , so 10 times smaller.

From (i) and (ii) it follows that for distinctiveness of *e* for *k*, the conditional probability of *e* given *k*,  $P(e|k)$ , counts for more than  $P(e|\neg k)$ . And this seems required. Consider, on the one hand, the uncommon feature ‘having 3 legs’. Although there are (presumably) relatively more dogs with three legs than there are other animals with three legs, this doesn’t mean that the generic ‘Dogs have three legs’ is true, or acceptable (cf. Leslie 2008). If a more common feature is used, on the other hand, an equally small difference between  $P(e|k)$  and  $P(e|\neg k)$  can make the difference between truth and falsity, or of (un)acceptability, of the generic sentence, if the generic is used to contrast *k* from other kinds.

In summary, the following analysis of generic sentences of the form ‘*ks* are *e*’ was proposed:

- ‘*ks* are *e*’ is true, or acceptable, if and only if *Repr*(*e*, *k*) is high.
- $Repr(e, k) =_{df} \Delta^{**}P_k^e \times Impact(e)$ .

It should be clear how examples like (4-a)–(4-c) can be accounted for on this proposal: (4-a) is true, or acceptable, because being striped is distinctive for tigers, whereas (4-b) is true because (i) more mosquitos than other types of insects carry the West Nile virus, and (ii) carrying this dangerous virus has a high impact. In van Rooij (2017), van Rooij and Schulz (in press) it is argued that a wide variety of generics can be accounted for using the above analysis, especially if (i) we make use of the context-dependence of which alternatives are relevant, and (ii) we assume that it is not just relative frequency that counts, but rather *stable* relative frequencies: it is not only that the measure  $P(e|k) - P(e|\neg k)$  should be high, but this measure should remain high when conditioned on relevant backgrounds.<sup>10</sup>

Moreover, we have argued that a high value of *Repr*(*e*, *k*) gives rise to the (perhaps false) impression that  $P(e|k)$  is high, thereby accounting for the general (but false) intuition that generics like ‘*ks* are *e*’ are true, or acceptable, just in case  $P(e|k)$  is high (if *P* measures frequencies). In van Rooij and Schulz (in press) we do this by making use of Tversky and Kahneman’s (1974) Heuristics and Biases approach. In van Rooij and Schulz (submitted), instead, we appeal to Pavlovian associative learning, for error- and competition-based learning formulas describing the learning process can converge in the long run to measures of

<sup>10</sup> The notion of stability is required to think of  $P(e|k) - P(e|\neg k)$  as helping to account for inductive generalizations, and does the work that Cohen (1999) argues his condition of ‘homogeneity’ should do. It is by concentrating ourselves on probabilities that are *stable* under conditionalization by various conditions that generics like ‘Bees are sterile’, ‘Israeli live along the coast’ and ‘People are over three years old’ are predicted to be bad, or false, although in each case the majority of the ‘kind’ has the relevant feature. For an analysis of stability that we favor, see Skyrms (1980).

distinctiveness as discussed above. It is well-known, for instance, that Rescorla and Wagner’s (1972) famous associative learning rule converges in the long run to the measure of contingency (cf. Chapman and Robbins 1990). More recently, Yuille (2006) has shown that a very similar learning rule converges to Shep’s measure of relative difference. Important for present purposes is that these learning rules not only describe the development of the associative strength between cue  $k$  and outcome  $e$ . They also are taken to measure the expectations of the learner to observe the outcome given a new encounter with the cue. Building on this idea, it is natural to propose that the subjective probability of a member of group  $k$  having feature  $e$  is given by how strong the agent expects any member of  $k$  to have feature  $e$ . It follows that subjectives probabilities can be very different from relative frequencies, because the former are based on distinctiveness.

One obvious objection to the above descriptive analysis in terms of (stable) frequencies should be mentioned, though:  $\Delta^*P_k^e$  by itself cannot account for the ‘intensional component’ of generic sentences showing in their ‘non-accidental’ understanding. Even if actually (by chance) all ten children of Mr. X are girls, the generic ‘Children of Mr. X are girls’ still seems false or inappropriate.<sup>11</sup> The sentence only seems appropriate if being a child of Mr. X somehow *explains why* one is a girl. In this paper we will explore to what extent we can explain the meaning of generic sentences in terms of *inherent dispositions* or *causal powers*. Even though such dispositions were philosophically suspect in much of the 20th century, we take such an exploration as a worthwhile enterprise, because it seems to be in accordance with many people’s intuition. Moreover, by adopting a causal stance, the non-accidental understanding of generics can, arguably, be explained as well.

### 3 Causal Readings of Generics

#### 3.1 Causal Explanation of Correlations

The theory of generics in terms of the measure  $\Delta^*P_k^e$  is very Humean, built on frequency data and probabilistic dependencies and the way we learn from those. Many linguists and philosophers feel that there must be something more: something hidden underlying these actual dependencies that **explains** them. A most natural explanation is a **causal** one: the probabilistic dependencies exists

<sup>11</sup> To account for such cases—i.e., the ‘unbounded’ character of generics—, Cohen (1999) makes use of limiting relative frequencies. The (causal) solution we will propose in the following sections will be different, but based on a similar intuition.

*in virtue of* objective kinds which have causal powers, capacities or dispositions.<sup>12</sup> Indeed, traditionally philosophers have assumed that the natural world is objectively divided into kinds, which have essences, a view that has gained popularity in the 20th century again due to the work of Kripke (1972/80) and Putnam (1975). A closely associated modern view that has gained popularity recently has it that causal powers (Harré and Madden 1975), capacities (Cartwright 1989) or dispositions (Shoemaker 1980; Bird 2007) are the truth-makers of laws and other generalities.<sup>13</sup>

Whereas probabilistic (in)dependencies are symmetric,<sup>14</sup> causal power relations are not. But neither are generic sentences. Such sentences of the form ‘ $k$ s are  $e$ ’ are, by their very nature, stated in an asymmetric way: first the noun  $k$ , then feature  $e$ . This naturally gives rise to the expectation that objects of type  $k$  are associated with features of type  $e$  because the former has the power to cause the latter. Where the goal of van Rooij and Schulz (in press) was to develop a semantic analysis of generic sentences that is descriptively adequate, the goal of this paper is to investigate to what extent this theory can be explained by basing it on an analysis of (perhaps unobservable) causal powers. In a sense, the answer to this question is quite clear: Shep’s notion of relative difference closely corresponds to Good’s (1961) measure of ‘causal support’:  $\log \frac{P(-e|-k)}{P(-e|k)}$ . In fact, Good’s notion is ordinally equivalent to Shep’s notion in the sense that  $\Delta^*P_k^e > \Delta^*P_{k^*}^{e^*}$  iff  $\log \frac{P(-e|-k)}{P(-e|k)} > \log \frac{P(-e^*|-k^*)}{P(-e^*|k^*)}$  for all  $e, e^*, k$  and  $k^*$ .<sup>15</sup> This is very interesting. In the end, though, also Good’s notion is just a frequency measure. What we would like to find is a ‘deeper’ foundation of our measure. In a sense, this is what Good provides as well, for he provides an axiomatization of his notion of causal support. But we think that the causal

<sup>12</sup> As noted already in the introduction, it seems no accident that (general) causal statements typically are of generic form: ‘Sparks cause fires’, ‘Asbestos causes cancer’.

<sup>13</sup> According to Strawson (1989), even Hume himself believed in causal powers.

<sup>14</sup>  $P(e|k) > P(e)$  iff  $P(k|e) > P(k)$  and  $P(e|k) = P(e)$  iff  $P(k|e) = P(k)$ .

<sup>15</sup> This was shown to one of the authors by Vincenzo Crupi. By a slight simplification, his proof comes down to the following:

$$\begin{aligned} \Delta^*P_k^e &= \frac{P(e|k) - P(e|-k)}{1 - P(e|-k)} = \frac{1 - P(-e|k) - 1 + P(-e|-k)}{P(-e|-k)} = \frac{P(-e|-k) - P(-e|k)}{P(-e|-k)} \\ &= \frac{P(-e|-k)}{P(-e|-k)} - \frac{P(-e|k)}{P(-e|-k)} = 1 - \frac{P(-e|k)}{P(-e|-k)}. \end{aligned}$$

Likewise,  $\Delta^*P_{k^*}^{e^*} = 1 - \frac{P(-e^*|k^*)}{P(-e^*|-k^*)}$ . Thus,

$$\begin{aligned} \Delta^*P_k^e > \Delta^*P_{k^*}^{e^*} &\text{ iff } 1 - \frac{P(-e|k)}{P(-e|-k)} > 1 - \frac{P(-e^*|k^*)}{P(-e^*|-k^*)} &\text{ iff } \frac{P(-e|k)}{P(-e|-k)} < \frac{P(-e^*|k^*)}{P(-e^*|-k^*)} \\ &\text{ iff } \frac{P(-e|-k)}{P(-e|k)} > \frac{P(-e^*|-k^*)}{P(-e^*|k^*)} &\text{ iff } \log \frac{P(-e|-k)}{P(-e|k)} > \log \frac{P(-e^*|-k^*)}{P(-e^*|k^*)}. \end{aligned}$$

foundation that we will give is more natural, and fundamental.

We don't want to claim that a causal analysis can account for all types of generics. Generics like 'People born in 1990 reach the age of 40 in the year 2030' and 'Bishops move diagonally' (in chess) are most naturally not treated in a causal way. Linguists (e.g., Lawler 1973; Greenberg 2003) also make a difference between generics formulated in terms of bare plurals (BPs) ('Dogs bark'), on the one hand, and generics stated in terms of indefinite singular (IS) noun phrases ('A dog barks'), and found that IS generics have a more limited felicity, and suggested that in contrast to a BP generic, for an IS generic to be felicitous, there has to exist a 'principled connection' between subject noun and predicate attributed to it. Perhaps this means that only IS generics should be given a causal analysis. Perhaps. But we do think that for many, if not most, BP generics causality could play an important role as well. The purpose of this paper is not to defend the strong view that all generics should be analyzed causally. Instead, our purpose is more modest: to explore the possibility of a causal power analysis of BP generics.<sup>16</sup>

As part of this, we want to clarify what, if any, advantage(s) such a causal power analysis might provide. These advantages could be of a conceptual and an empirical nature. As for the former, if all that is gained by a causal analysis of e.g., 'Aspirin relieves headaches' is that the observed frequency of relieved headaches is said to be due to the Aspirins' unobservable capacity to relieve headache, nothing is won. For a causal analysis to be useful more insights should be gained, for instance in the internal structure of the cause. But a causal analysis can be useful here as well, as shown by the recent abundance of papers on *mediation* (e.g. Preacher and Kelley 2011; Pearl 2014): causal models can (be used to) explain not only *why* something happened, but also *how* it happened. Scientists are not only interested to learn *that* Aspirin relieves headaches, they are also interested in the *mechanism* by which it does so. Although in this paper we won't make use of the recent insights of causal mediation analyses that make a difference between *direct* and *indirect* causal effects, we think that this can be useful for the analysis of generics involving social kinds as well. In the next section we will show that under certain circumstances a causal interpretation gives rise to different, and arguably *more adequate predictions* than an extensional theory making use of  $\Delta^{**}P_k^e$ . But first we will show in this section that under natural assumptions a causal analysis *explains* the predictions made by using  $\Delta^{**}P_k^e$ .

<sup>16</sup> We leave a discussion of IS generics to another paper.

### 3.2 A Causal Derivation of $\Delta^{**}P_k^e$

For our causal explanation of the measure  $\Delta^{**}P_k^e$  we follow Cheng (1997) and assume that objects of type  $k$  have unobservable **causal powers** to produce features of type  $e$ . We will denote this unobservable causal power by  $p_{ke}$ . It is the probability with which  $k$  produces  $e$  when  $k$  is present in the absence of any alternative cause. This is different from  $P(ek)$ . The latter is the relative frequency of  $e$  in the presence of  $k$ . We will denote by  $u$  the (unobserved) alternative potential cause of  $e$  (or perhaps the union of alternative potential causes of  $e$ ), and by  $p_{ue}$  and  $P(elu)$  the causal power of  $u$  to produce  $e$  and the conditional probability of  $e$  given  $u$ , respectively. We will assume (i) that  $e$  does not occur without a cause and that  $k$  and  $u$  are the only potential causes of  $e$  (or better that  $u$  is the union of all other potential causes of  $e$  other than  $k$ ), i.e., that  $P(e|-k, -u) = 0$ , (ii) that  $p_{ke}$  is independent of  $p_{ue}$ , and (iii) that  $p_{ke}$  and  $p_{ue}$  are independent of  $P(k)$  and  $P(u)$ , respectively, where independence of  $p_{ke}$  on  $P(k)$  means that the probability that  $k$  occurs and produces  $e$  is the same as  $P(k) \times p_{ke}$ . The latter independence assumptions are crucial: by making them we can explain the *stability* and (relative) context-independence of generic statements.

Now we are going to derive  $p_{ke}$ , the causal power of  $k$  to produce  $e$ , following Cheng (1997).<sup>17</sup> To do so, we will first define  $P(e)$  assuming that  $e$  does not occur without a cause and that there are only two potential causes,  $k$  and  $u$ , i.e.,  $P(e|-k, -u) = 0$  (recall that  $P(k \vee u) = P(k) + P(u) - P(k \wedge u)$ ):

$$(6) \quad P(e) = P(k) \times p_{ke} + P(u) \times p_{ue} - P(k \wedge u) \times p_{ke} \times p_{ue}.$$

In case of a *controlled experiment*, we can set (and not just observe)  $u$  to be false. In that case  $p_{ke}$  is nothing else but the probability of  $e$ , conditional on  $k$  and  $-u$ :

$$(7) \quad p_{ke} = P(e|k, -u) \quad \text{the causal power of } k \text{ to generate } e.$$

One problem with this notion is that controlled experiments are hard, especially if we don't know really what this union of alternative causes  $u$  is. Thus, it still remains mysterious how anyone could know, or reasonably estimate, the causal power of  $k$  to produce  $e$ . It turns out that we can still measure this causal power even if we don't know exactly what  $u$  is, if we **assume that  $k$  and  $u$  are, or are believed to be, independent of each other**. Assuming independence of  $k$  and  $u$ ,  $P(e)$  becomes

<sup>17</sup> Glymour (2001, p. 75) calls Cheng's derivation 'a brilliant piece of mathematical metaphysics'. Cheng (1997) also discusses an analysis of *preventive* causes. We won't deal with this in this paper.

$$(8) \quad P(e) = P(k) \times p_{ke} + P(u) \times p_{ue} - P(k) \times P(u) \times p_{ke} \times p_{ue}.$$

As in Sect. 2,  $\Delta P_k^e$  is going to be defined in terms of conditional probabilities:

$$(9) \quad \Delta P_k^e = P(e|k) - P(e|\neg k).$$

The relevant conditional probabilities are now derived as follows (by changing  $P(\cdot)$  in (8) into  $P(\cdot|k)$  or  $P(\cdot|\neg k)$ ):

$$(10) \quad \begin{aligned} P(e|k) &= p_{ke} + (P(u|k) \times p_{ue}) - p_{ke} \times P(u|k) \times p_{ue} \\ P(e|\neg k) &= P(u|\neg k) \times p_{ue} \quad (\text{derived from (8), because } P(k|\neg k) = 0). \end{aligned}$$

As a result,  $\Delta P_k^e$  comes down to

$$(11) \quad \begin{aligned} \Delta P_k^e &= p_{ke} + (P(u|k) \times p_{ue}) - (p_{ke} \times P(u|k) \times p_{ue}) - (P(u|\neg k) \times p_{ue}) \\ &= [1 - (P(u|k) \times p_{ue})] \times p_{ke} + [P(u|k) - P(u|\neg k)] \times p_{ue}. \end{aligned}$$

From this last formula we can derive  $p_{ke}$  as follows:

$$(12) \quad p_{ke} = \frac{\Delta P_k^e - [P(u|k) - P(u|\neg k)] \times p_{ue}}{1 - P(u|k) \times p_{ue}}.$$

From (12) we can see that  $\Delta P_k^e$  gives a good approximation of causal power in case (i)  $u$  is independent of  $k$  (meaning that  $P(u|k) - P(u|\neg k) = 0$ ), and (ii)  $p_{ue} \times P(u|k)$  is low. Obviously, in case  $k$  is the only potential direct cause of  $e$ , i.e., when  $p_{ue} = 0$ , it holds that  $p_{ke} = \Delta P_k^e$ . Because in those cases  $P(e|\neg k) = 0$ , it even follows that  $p_{ke} = P(e|k)$ .

Our above derivation shows that to determine  $p_{ke}$  in case events or features of type  $e$  might have more causes, we have to know the causal power of  $p_{ue}$ , which is equally unobservable as  $p_{ke}$ . You might wonder what we have learned from the above derivation for such circumstances. It turns out, however, that  $p_{ke}$  can be estimated in terms of observable frequencies after all, because we assumed that  $P(k)$  and  $P(u)$  are independent of each other. On this assumption it follows that  $P(u|k) = P(u) = P(u|\neg k)$  and that (12) comes down to

$$(13) \quad p_{ke} = \frac{\Delta P_k^e}{1 - P(u|k) \times p_{ue}}.$$

Because of our latter independence assumption, it follows as well that  $P(u|k) \times p_{ue} = P(u) \times p_{ue} = P(e|\neg k)$ . This is because  $P(u) \times p_{ue}$  is the probability that  $e$  occurs and is produced by  $u$ . Now,  $P(e|\neg k)$  estimates  $P(u) \times p_{ue}$  because  $k$  occurs independently of  $u$ , and, in the absence of  $k$ , only  $u$  produces  $e$ . It follows that  $p_{ke}$  can be defined in terms of observable frequencies as follows:

$$(14) \quad p_{ke} = \frac{\Delta P_k^e}{1 - P(e|\neg k)} = \frac{P(e|k) - P(e|\neg k)}{1 - P(e|\neg k)}.$$

But this is exactly the same as  $\Delta^* P_k^e$ , the measure in terms of which we have stated the truth, or acceptability, conditions of generic sentences in Sect. 2! Thus, in case we assume that a generic sentence of the form ‘Objects of type  $k$  have feature  $e$ ’ is true, or acceptable, because objects of type  $k$  cause, or produce, features of type  $e$ , we derive exactly the semantics we have proposed in the first place (if  $\alpha = \frac{1}{2}$ ). It follows that as far as our descriptive analysis of generics in Sect. 2 in terms of  $\Delta^* P_k^e$  was correct, what we have provided in this section is a **causal explanation**, or grounding, of this descriptive analysis.

Let us go back to the case that we talk about a **controlled experiment** where we set the alternative causes,  $u$ , to 0. Thus, for this controlled experiment we only have to look at the probability function conditioned by  $\neg u$ , i.e.,  $P(\cdot|\neg u)$ . Because we know by assumption that  $P(e|\neg k, \neg u) = 0$ , it immediately follows that

$$p_{ke}^{\neg u} = \frac{P(e|k, \neg u) - P(e|\neg k, \neg u)}{1 - P(e|\neg k, \neg u)} = P(e|k, \neg u).$$

Thus, for the controlled experiment where we set  $u$  to be false, we see that the causal power of  $k$  to generate  $e$  is just  $P(e|k, \neg u)$ , just as we claimed before.

The above derivation of  $p_{ke}$  causally motivated Shep’s notion of ‘relative difference’. But that notion is a special case of  $\Delta^{**} P_k^e$  in case  $\alpha = \frac{1}{2}$ . We have seen above that in case  $\alpha = 1$ , what should come out is that  $\Delta^{**} P_k^e$  comes down to  $P(ek)$ . Does a causal analysis motivate this as well? It does! To see this, notice that in case  $k$  is the only potential cause of  $e$ , it immediately follows from (6) that  $P(e)$  can be determined as follows:

$$(15) \quad P(e) = P(k) \times p_{ke}.$$

As a result,  $P(ek)$  reduces to  $p_{ke}$ . Thus,  $p_{ke} = P(e|k)$  in case  $k$  is the only potential cause of  $e$ , just like  $\Delta^{**} P_k^e$  came down to  $P(ek)$  in case  $Alt(k) = \emptyset$ . We conclude that our earlier measure  $\Delta^{**} P_k^e$  could be motivated by our causal powers view both when  $\alpha = \frac{1}{2}$  and when  $\alpha = 1$ .

How do these causal powers account for generic sentences? This is easiest to see for generics involving homogenous substances, like ‘Sugar dissolves in water’ and ‘metal conducts electricity’. Intuitively, these are true, because of the causal power of sugar and metal to generate the observable manifestations that come with the relevant predicates. Similarly, ‘Tigers are striped’ is true, on a causal account, because of what it is to be a tiger. But sometimes the power description should be relativized. For instance, ‘Ducks lay eggs’ is true, although only the female chicken do so. Intuitively, it is not the causal power of ‘being a duck’ in general that makes this generic true. Rather, it is the causal power of being a *female* duck. But this comes out naturally. Cohen (1999) argued that the

‘domain’ of the probability function should be limited to individuals that make at least one of the natural alternatives of the predicate term true. In our example, it is natural to assume that  $Alt(lay\ eggs) = \{Lay\ eggs, give\ birth\ live\}$ . Because  $\bigcup Alt(lay\ eggs) \approx Female$ , this means that we should only consider female ducks. This should be done as well for the estimation of causal power. Doing so, it will be the case that the causal power of female ducks to lay eggs is high, which gives rise to the correct prediction that the generic ‘Ducks lay eggs’ is true. It is also clear how our analysis can account for ‘striking’ generics like (4-b) and (4-c): instead of demanding that  $\Delta^*P_k^e \times Impact(e)$  is high, one now demands that  $p_{ke} \times Impact(e)$  is high, which normally comes down to the same.

In the derivation above we have assumed that  $k$  by itself can cause  $e$ . Of course, this is a simplification. Striking a match, for instance, does not by itself cause it to light. Certain background conditions have to be in place: there must be oxygen in the environment, the match must be dry, etc. In a sense this is captured: we don’t assume that  $p_{ke}$ , or  $\Delta^*P_k^e$ , is either 1 or 0. In fact, we can think of  $\Delta^*P_k^e$  as modeling the probability with which the background conditions are in place (Cheng 2000). To see this more precisely, let us follow Cheng and Novick (2004) and be more explicit about this by taking *background causes* more explicitly into account. Suppose that  $k$  can interact with  $i$  to cause  $e$ . Let us also assume that just like  $k$ ,  $u$  and the interaction  $ki$  are generative cause, and not preventive ones.<sup>18</sup> Notice that given independence,  $P(e)$  is now the complement of the chance that  $e$  is failed to be generated by any of the three causes:

$$(16) \quad P(e) = 1 - [1 - P(k) \times p_{ke}] \times [1 - P(u) \times p_{ue}] \times [1 - P(k) \times P(i) \times p_{ki,e}].$$

Thus, assuming independence,

$$(17) \quad \begin{aligned} \text{a. } P(e|\neg k) &= P(u) \times p_{ue} \text{ and} \\ \text{b. } P(e|k) &= 1 - [1 - p_{ke}] \times [1 - P(u) \times p_{ue}] \times [1 - P(i) \times p_{ki,e}]. \end{aligned}$$

Subtracting (17-a) from (17-b) gives us

$$(18) \quad \begin{aligned} \Delta P_k^e &= p_{ke} + P(i) \times p_{ki,e} - P(e|\neg k) \times p_{ke} \\ &\quad - P(e|\neg k) \times P(i) \times p_{ki,e} \\ &\quad - P(i) \times p_{ke} \times p_{kie} + P(e|\neg k) \times P(i) \\ &\quad \times p_{ke} \times p_{ki,e}. \end{aligned}$$

But this means that

$$(19) \quad \Delta P_k^e = [p_{ke} + P(i) \times p_{ki,e} - P(i) \times p_{ke} \times p_{ki,e}] \times [1 - P(e|\neg k)].$$

<sup>18</sup> Cheng (2000), Cheng and Novick (2004) give also derivations when some of these are preventive causes.

Rearranging things gives us

$$(20) \quad \Delta^*P_k^e = \frac{\Delta P_k^e}{1 - P(e|\neg k)} = p_{ke} + P(i) \times p_{ki,e} - P(i) \times p_{ke} \times p_{ki,e}.$$

In case we know that  $p_{ke} = 0$ , as in the case of the match and the oxygen,

$$(21) \quad \Delta^*P_k^e = P(i) \times p_{ki,e}.$$

Thus for predicting the lighting of the match when it is struck  $\Delta^*P_k^e$  is still useful, because it measures the causal power of  $k$  to produce  $e$ , given background conditions  $i$  (oxygen, dryness of the surrounding air). If the background conditions for  $k$  to produce  $e$  are stable (say  $P(i) = 1$ ), then  $p_{ki,e} = \Delta^*P_k^e$ . Finally, in case  $p_{ke} = 0$  and  $p_{ki,e} = 1$ , the measure  $\Delta^*P_k^e$  estimates  $P(i)$ , the probability with which the background conditions are in place. We think that in all these cases, if  $\Delta^*P_k^e$  is high, the corresponding generic is considered true, or acceptable.

What if the conjunctive cause  $k \wedge i$  is the *only* potential cause of  $e$ ? One can easily see that in that case  $\Delta^{**}P_k^e = \Delta^*P_k^e = \Delta P_k^e = P(e|k)$ . It is also easy to see that now  $P(e|k) = P(i) \times p_{ki,e}$ , and thus that also  $\Delta^*P_k^e = P(i) \times p_{ki,e}$ .

The result of this section that  $p_{ke}$  can be estimated by the observable measure  $\Delta^{**}P_k^e$  was partly due to our assumption that  $k$  is probabilistically independent of alternative causes for  $e$ . In the following section we will investigate what the relation between the two measures  $p_{ke}$  and  $\Delta^{**}P_k^e$  will be when we give up this independence assumption. But notice that in this section we also saw that  $\Delta^{**}P_k^e$  is also a good measure of the causal power of  $k$  to produce  $e$  even if  $k$  can produce  $e$  only given background condition  $i$ . In that case it measures  $P(i) \times p_{ki,e}$ . But also in this derivation independence assumptions are made, and it is interesting to see as well what happens if we give up these independence conditions used in that derivation.

Before we will give up on the above independence assumption, let us first suggest how our causal powers can be used not only for generics, but for other types of sentences as well.

### 3.3 Habitual Sentences and Disposition Ascriptions

Until now we have discussed generic sentences, sentences that involve kinds, or groups of individuals. But some sentences just involving one object, or individual, behave semantically in a very similar way. In linguistics, a distinction is made between *episodic* sentences and *habitual* ones. Episodic sentences are about particular times, places and events, but habitual sentences are not. Sentences in the simple past like (22-a) are typically episodic sentences, while habitual

sentences like (22-b)–(22-d) (like generic sentences involving groups), are typically stated using the present tense.

- (22) a. John drank milk with lunch today.  
 b. Johns drinks milk with lunch.  
 c. Mary smokes  
 d. Sue works at the university.

Just like generic sentences, also habitual sentences express generalities that typically allow for exceptions. Within semantics, generics and habituals are normally treated similarly (e.g. Krifka et al. 1995). Habituals differ from generics only because the generalizations they express do not involve multiple individuals, but rather multiple events involving a single individual. Although some habitual sentences are true, or acceptable, just because of high (stable) conditional probability (perhaps e.g. (22-b)), it makes even less sense for most habituals than for most generics to assume that their truth, or acceptability, conditions always demand high conditional probability, or normality, with respect to the events the relevant individual is involved. This is already clear for examples like (22-c)–(22-d), but is immediately obvious for the habitual reading of an example like

- (23) Paul picks his nose.

For this sentence to be true, or acceptable, we don't demand that Paul normally, or most of the time, is picking his nose. Moreover, just like for generics, it seems that *impact* plays a major role. As observed already by Carlson (1977), it takes much less killing-events involving Mary to make the habitual (24-a) true, or acceptable, than smoking-events involving her to make (22-c) true, or acceptable.

- (24) a. Mary murders children.  
 b. Hillary Clinton is a liar

The reason is the *impact* of murdering children, or so we assume. Trump's successful rhetorical use of the habitual (24-b) in the 2016 USA-presidential election campaign (where the issue was whether Clinton lied about important classified information) only corroborates this. All this suggests that from a descriptive point of view, habituals should be treated like generics, demanding high  $\Delta^{**}P_k^e \times \text{Impact}(e)$  for its truth or acceptability.

But just like for generics, this frequency-like analysis leaves open the **explanatory reason** why. Moreover, a frequency-based analysis cannot explain the **intensional** character of habitual sentences.<sup>19</sup> Suppose that Sue's function is to handle the mail from Antarctica, although no

mail ever came from there yet. Then the habitual (25) is, intuitively, still true.

- (25) Sue handles the mail from Antarctica.  
 (from Krifka et al. 1995)

This suggests that a causal power analysis of habitual sentences—demanding that  $p_{ke} \times \text{Impact}(e)$  rather than  $\Delta^{**}P_k^e \times \text{Impact}(e)$ , is high—is natural. But what should variable  $k$  now denote? Intuitively, it should be something like the individual's *character*, *personality*, *temperament*, or (sometimes) *function*. Thus, on a causal power analysis, habituals like (22-b)–(22-d) are taken to be true *due to* something inherent of John, Mary and Sue, respectively. Such a causal power analysis of habituals will no doubt be controversial, but we do believe that habituals like (23), (24-a) and (24-b) have their societal effect exactly *because* we read habituals this way: these sentences say something about the (stable) *characters* of the individuals involved! Similarly, it seems natural to use causal powers for the analysis of what linguists call 'individual-level' predicates like 'being intelligent' and 'being blond'. Such predicates are contrasted with so-called 'stage'-level predicates, and the difference is that only the former are taken to be stable over time, and that sentences in which they are used say something about the character or disposition of the person(s) they are predicated of. Indeed, Chierchia (1995) proposed already that individual-level predicates are inherent generics.

The distinction between episodic and non-episodic sentences occurs also for other types of sentences:

- (26) a. This sugar lump is dissolving in water now.  
 b. This sugar lump dissolves in water.

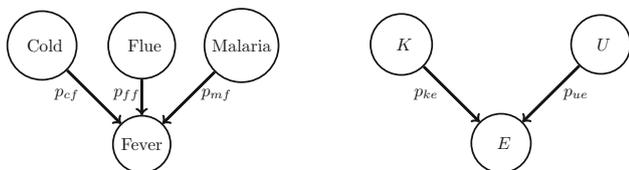
Whereas (26-a) describes the occurrence of an event, (26-b) describes, intuitively, a *dispositional* property of an object. Within analytic philosophy, two analyses of dispositional sentences have been widely discussed: a conditional one, favored by Ryle (1949), Goodman (1954), and a kind-based analysis suggested by Quine (1970). In van Rooij and Schulz (to appear) we argued in favor of a causal analysis of Quine's suggestion: this lump is of the kind sugar and it dissolves in water *because* sugar has the causal power to dissolve in water. We argue that this analysis overcomes many problems of alternative treatments of disposition ascriptions, and that the analysis is much less mysterious than it might look at first.

## 4 Giving Up Independence of the Potential Causes

In the previous section we assumed with Cheng (1997) that  $e$  had two potential causes,  $k$  and  $u$ , and that these causes were independent of each other:  $P(u|k) = P(u|\neg k) = P(u)$ .

<sup>19</sup> A purely frequency-based analysis also cannot account for the intuition that whereas 'Mary **usually** drinks a beer' is ok, 'Mary drinks a beer' does not have an habitual reading. Whether this can or should be explained by a causal power analysis we don't know.

As noted by Glymour (2001), by adopting this assumption, Cheng assumed implicitly a specific type of causal structure: that what via Pearl (1988, p. 184) is known as a ‘Noisy-OR gate’. Pearl (1988) introduced noisy OR-gates mainly for complexity reasons: it simplifies the calculation of, in our case,  $P(e)$ . To illustrate, consider a simple case. John has a fever. We want to explain why. What was the cause of his fever? There are several alternative hypotheses: it could be (let’s say) a cold, the flue, or malaria that caused his fever. If we don’t assume that the potential causes are independent of each other, it is very complex to determine the probability of getting a fever. With the independence assumption, however, things are much simpler. We can illustrate our case graphically by the following Noisy-OR gate to the left, where  $p_{cf}$ , for instance, denotes the causal power of a cold to induce fever. What Cheng (1997) uses is the picture on the right, which is of the same type.



In general, it can be very hard to determine  $P(Fever)$  given the probabilities of a set of potential causes. This changes if we assume independence. Now  $P(Fever)$  can be calculated as the complement of the chance that *Fever* is failed to be generated by any of the three causes. More generally, if  $k_1, \dots, k_n$  are the potential causes of  $e$ ,  $P(e|k_1, \dots, k_n)$  can now be calculated as follows:

$$(27) \quad P(e|k_1, \dots, k_n) = 1 - \prod_{k=1}^{i=n} (1 - p_{ke}).$$

This is exactly the way Pearl (1988) and others determine the probability of  $e$  given that the potential causes form a Noisy-OR gate.<sup>20</sup> And from this formula it immediately follows that  $p_{k_1e} = P(e|k_1, \neg u)$ , if  $u = k_2 \vee \dots \vee k_n$ , i.e., what is the causal power in a controlled experiment.

Thus, as noted by Glymour (2001), the models that Cheng uses to calculate how we can estimate causal powers are in fact special cases of structural causal models as developed by Pearl (2000), Spirtes et al. (2000). In general, the potential causes of a variable don’t have to be

independent of each other. Glymour (2001)<sup>21</sup> shows that also in such situations, the causal power of  $k$  to influence  $e$  can sometimes be estimated from frequency data, at least if we keep in mind the causal structure that generated these data.

If independence is only a useful, but sometimes incorrect, heuristics to determine probabilities, it raises the question what happens if we give up this independence assumption? Quantitatively speaking, there are two possibilities:  $P(u|k) > P(u|\neg k)$  and  $P(u|k) < P(u|\neg k)$ . Already by looking at the general definition of  $p_{ke}$ :

$$(28) \quad p_{ke} = \frac{\Delta P_k^e - [P(u|k) - P(u|\neg k)] \times p_{ue}}{1 - P(u|k) \times p_{ue}},$$

we can immediately observe the following:

1. If  $P(u|k) < P(u|\neg k)$ , then  $\Delta^* P_k^e$  **underestimates**  $p_{ke}$ .
2. If  $P(u|k) > P(u|\neg k)$ , then  $\Delta^* P_k^e$  **overestimates**  $p_{ke}$ .

Thus, although giving up on independence doesn’t allow us anymore to determine  $p_{ke}$  in terms of observed frequencies alone (because we now also need to know  $p_{ue}, P(u|k)$  and  $P(u|\neg k)$ ), giving up independence still potentially gives rise to interesting empirical consequences. In the following subsections we will look at both cases, and see that they give rise to interesting new predictions.

### 4.1 $\Delta^* P_k^e$ (Assuming Independence) Underestimates $p_{ke}$

First, we will look at the most extreme case where  $P(u|k) < P(u|\neg k)$ , namely where  $u$  and  $k$  are **incompatible**. Notice that in that case  $P(u|k) = 0$ . The relevant conditional probabilities are then derived from (6) as follows:  $P(e) = P(k) \times p_{ke} + P(u) \times p_{ue}$ . From this we derive immediately that  $P(e|k) = p_{ke}$ , because  $P(u|k) = 0$ . Notice that if we assume that  $k$  only produces  $e$  given background  $i$ , a similar observation shows that now  $P(e|k) = P(i) \times p_{ki,e}$ , if background condition  $i$  is independent of  $k$ .

Thus, we see that in case  $k$  and  $u$  are incompatible, the causal power of  $k$  to produce  $e$  is the same as the conditional probability  $P(e|k)$ , just as was the case if  $k$  is the only cause of  $e$ . Perhaps this can explain the intuition people have that the acceptability of a generic sentence of the form

<sup>20</sup> With Cheng we assumed that the potential causes are either ON or OFF. But for many potential causes there is no way to be OFF. Consider height, or weight of persons, for instance. If these are alternative causes of  $k$ , it doesn’t make sense to determine  $p_{ke}$  as  $P(e|k, \neg u)$ . If  $U$  is the variable ranging over different values it can take, the natural alternative is to use the formula  $\sum_{u \in U} [P(e|k, u) \times P(u)]$  (cf. Pearl 2000; Spirtes et al. 2000). For a generalization that allows also  $k$  to be neither ON nor OFF, see Danks (manuscript).

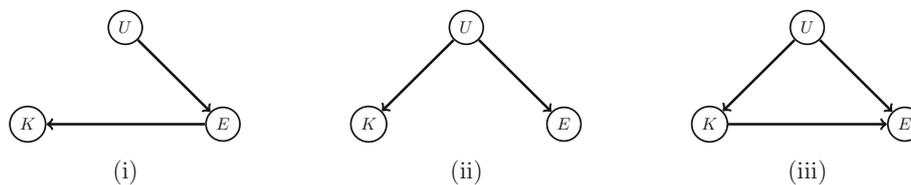
<sup>21</sup> Pearl (2000) has an alternative derivation of  $\Delta^* P_k^e$ , what he calls ‘the probability of causal sufficiency’, PS. Pearl derives PS—our  $\Delta^* P_k^e$ —from the measure  $P(e_k|\neg k, \neg e)$ , the probability of  $e$  after intervention with  $k$  when you are in a state where  $k$  and  $e$  are false. In this derivation Pearl doesn’t use the assumption that there is a statistically independent alternative cause,  $u$ , that may produce  $e$ . He substitutes this assumption with an assumption of monotonicity: that  $k$  never prevents  $e$ . Pearl doesn’t make use of causal powers in the derivation of  $\Delta^* P_k^e$ , but uses an assumption of causality, or intervention, as primitive instead. One can think of Pearl’s PS as a generalization of Cheng’s causal power as well, because applicable in more situations than Cheng’s notion.

'ks are  $e$ ' goes with its conditional probability  $P(ek)$ . Thus, although under natural independence conditions  $p_{ke} = \Delta^*P_k^e$ , this is no longer the case once  $k$  and  $u$  are not taken to be probabilistically independent.

Of course, one might take a causal view at  $\Delta^*P_k^e$ , or better, perhaps, a perspective on  $\Delta^*P_k^e$  where one doesn't assume that the potential causes of  $e$  are independent. We have seen in Sect. 3.2 that in case of a controlled experiment where we set  $u$  to 0, we can look at

$$\Delta^*P_{k,-u}^e = \frac{P(e|k, \neg u) - P(e|\neg k, \neg u)}{1 - P(e|\neg k, \neg u)} = P(e|k, \neg u)$$

. If we assume that  $k$  and  $u$  are incompatible, this reduces to  $P(ek)$ .



Are there good examples of generic statements where  $k$  and  $u$  (the union of alternative causes of feature  $e$ ) are incompatible, or where  $k$  is taken to be the only cause of  $e$ ? This depends very much on what one takes the alternative causes to be. Take any generic of the form 'ks are  $e$ '. Let us assume that  $P(ek)$  is high. We have argued in Sect. 2 that this is not always enough to make the generic true. But now suppose that ' $k$ ' denotes a kind of animal (e.g., 'horse') and that  $e$  is a feature like 'having a heart'. If one makes the Aristotelian assumption that  $x$  is a member of a kind if and only if  $x$  has the **essence** of that kind, then it is natural that we take the alternative causes of (having feature)  $e$  to be (essences of) other kinds of animals. Thus,  $u = \bigcup Alt(k)$ , with  $k$  incompatible with  $u$ . If for the analysis of generics we adopt the frequency measure  $\Delta^*P_k^e$  (with  $k$  denoting horses and  $e$  denoting creatures with a heart), the generic 'Horses have a heart' is most likely counted as false, or unacceptable, simply because  $P(e|k) = P(e|\neg k) = P(e|\bigcup Alt(k))$ , and thus  $P(e|k) - P(e|\neg k) = 0$ , meaning that also  $\Delta^*P_k^e = 0 = \Delta^*P_k^e$ , if  $\alpha = \frac{1}{2}$ . Thus, on a correlation-based analysis, the generic is predicted to be false if  $\alpha = \frac{1}{2}$ .<sup>22</sup> On a causal power view, however, the sentence is predicted to be true, because now  $p_{ke} = P(e|k) \approx 1$ . Of course, that  $p_{ke} = P(e|k)$  was due to the assumption that  $k$  and  $u$  (the union of alternative causes of feature  $e$ ) are incompatible, a view that makes perhaps sense only once one makes the highly controversial Aristotelian assumption that it is the essence of a kind that has causal powers.

<sup>22</sup> Of course,  $\Delta^*P_k^e$  comes down to  $P(ek)$ , if  $\alpha = 1$ .

Controversial as this assumption might be, psychologists like Keil (1989), Gelman (2003) and others have argued that both children and adults tend to have essentialist beliefs about a substantial number of categories, and in particular about natural kinds like water, birds and tigers.<sup>23</sup>

#### 4.2 $\Delta^*P_k^e$ Overestimates $p_{ke}$ : Some Challenges

The causal power of  $k$  to produce  $e$ ,  $p_{ke}$ , will be lower than  $\Delta^*P_k^e$  in the following three causal structures,<sup>24</sup> because in these structures there is either no causal relation from  $k$  to  $e$ , or  $u$  is a confounding factor to determine the causal influence of  $k$  on  $e$  in terms of conditional probabilities (iii) (but also (ii)):

Intuitively, in cases (i) and (ii) it should be that although  $P(ek)$  can be high, still  $k$  doesn't have any causal power to produce  $e$ , i.e.,  $p_{ke} = 0$ . Indeed, this is what comes out. To see this for (i), recall that we noted in Sect. 3 that in a controlled experiment  $p_{ke}$  comes down to  $P(e|k, \neg u)$ , where  $u$  denotes the disjunction of all potential causes of  $e$  different from  $k$ . But it is obvious that for (i) this means that  $p_{ke} = P(e|k, \neg u) = 0$ , because now there is nothing that could cause  $e$ .<sup>25</sup> In the picture in the middle,  $u$  is a

<sup>23</sup> Danks (2014) represents concepts as graphical-model-based probability distributions (see also Rehder 2003; Sloman 2005). He shows that all the most prominent models of concepts (the theory-based, the prototype-based, and the exemplar-based) can be modeled by such distributions. An exemplar-based model of a concept, for instance, according to which the connection between an individual  $d$  and a concept  $C$  should be based on the similarity between  $d$  and each of the exemplars of  $C$ , can be represented by a probability function over features, such that all pairs of features are associated with one another, but that these associations are all due to an *unobserved common cause*. (Danks 2014 shows how to directly translate in both directions between an exemplar-based concepts—making use of similarities between the members—and a graphical-based probability function with a common cause structure.) Arguably, this is just as well the correct representation of a probabilistic version of a more traditional essence-based model of concepts, with the essence, or substantial form, as the unobserved, or latent, variable.

<sup>24</sup> To be sure, there are much more complicated causal structures with many more variables where this will be the case. To focus discussion, however, we look only at these simple cases.

<sup>25</sup> Alternatively, we might follow Pearl (2000) and measure the relevant causal power in terms of intervention as follows  $P(k_e|\neg k, \neg e)$ . But because in causal structure (i) intervention of  $k$  doesn't influence the probability of  $e$ , and because now  $\neg e$  is taken to be true, this means that  $p_{ke} = P(e_k|\neg k, \neg e) = 0$ , just as it should be.

**common cause** of  $k$  and  $e$ , and also now  $u$  is the only cause of  $e$  and as a result  $P(e)$  is just  $P(u) \times p_{ue}$ .<sup>26</sup> Although  $p_{ke} = 0$  in causal structures (i) and (ii), it is clear that there are examples of the form ‘ $ks$  are  $e$ ’ with these causal structures that are intuitively true, or acceptable, perhaps because  $\Delta^*P_k^e$  is high.

Most obviously problematic for the causal analysis we have presented so far are acceptable generics of the form ‘ $ks$  are  $e$ ’ with causal structure (i). That such examples exist can easily be shown, for both of the following two generics seem true, or acceptable:

- (29) a. People that are nervous smoke.
- b. People that smoke are nervous.

It is obvious that one cannot account for the truth, or acceptability, of *both* examples by saying that the subject-term causes the predicate to hold. So, what can a causal analysis say about these examples? That seems a serious challenge.

A well-known example of common cause structure (ii) involves yellow fingers ( $k$ ) and lung cancer ( $e$ ). It used to be the case that cigarettes had filters that caused smokers to get yellow fingers. We know by now that smoking also causes lung cancer. It follows that many people that have yellow fingers get lung cancer, and thus that  $\Delta^*P_k^e$  (and  $P(ek)$ ) is high. But, obviously, getting lung cancer is not due to having yellow fingers, i.e., in this causal structure  $p_{ke} = 0$ . It is smoking ( $u$ ) that causes both. However, the following generic is arguably still true, or acceptable:

- (30) People with yellow fingers develop lung cancer.

We are less sure whether acceptable generics of the form ‘ $ks$  are  $e$ ’ exist for structure (iii), though we will discuss a potential counter-example involving this structure as well. Suppose that women drink significantly more tea on a regular basis than men and that it is somewhat better to drink tea than to drink, say, coffee. In many countries it is also the case that women have a higher life expectancy than the average life expectancy. Thus, there will be a positive correlation between ‘drinking tea’ and ‘higher than

average life expectancy’. We wonder whether this by itself makes the following generic true.

- (31) People that drink tea regularly have a higher than average life expectancy.

If this generic is taken to be true, or acceptable, it again poses a challenge to the causal analysis pursued until now. With one of the reviewers of this paper, we have serious doubt about the truth, or acceptability, of (30), and therefore leave the discussion of generics in causal structures (iii) for what they are in this paper.

## 5 Towards a more General Causal Analysis

Until now we have assumed that on a causal analysis of generics, ‘ $ks$  are  $e$ ’ is true, or acceptable, if and only if  $p_{ke}$  is high. Some examples in the previous section show clear counterexamples to that: a high  $p_{ke}$  might be *sufficient* condition for the generic to be true, or acceptable, it is certainly *not* a *neccessary* condition. This holds in particular for causal structure (i) above, where  $e$  is a cause for  $k$ . Indeed, the most obvious predicted difference between the associative analysis based on  $\Delta^*P_k^e$  and the causal analysis based on  $p_{ke}$  is that the latter causal analysis is essentially *asymmetry*, while the former correlation-based analysis need not be. This is similar to causal versus non-causal analyses of counterfactuals. Whereas Lewis’ (1973b) similarity-based analysis of counterfactuals is not necessarily asymmetric, more recent causal analyses that follow Pearl (2000) are. As a result, these causal analyses have a problem to explain how to account for so-called ‘back-tracking counterfactuals’ like ‘If she came out laughing, her interview went well’, counterfactuals in which the consequent cannot have been caused by the antecedent because the latter came later in time than the former.

Suppose we have a causal structure of the form  $k \rightarrow e \leftarrow u$ . It is well possible that in such cases  $\Delta^*P_k^e = \frac{P(k|e) - P(k|\neg e)}{1 - P(k|\neg e)}$  has a high value, meaning that generics of the form ‘Objects of type  $e$  are (generally) of type  $k$ ’ are true in such circumstances according to the non-causal analysis discussed in Sect. 2. On the causal analysis presented above, however,  $p_{ek} = 0$ , as we saw. But how, then, can we account for the truth, or acceptability, of *both* (29-a) and (29-b)?

Perhaps such examples simply show that causality is not *semantically* relevant for the analysis of generics, it is at most relevant for *pragmatics*: people take, perhaps wrongly, generics to say something about causal powers. Perhaps. But even then we would need a causal analysis for (29-a) and (29-b) within pragmatics. We believe that we can provide a causal analysis for both types of generics.

<sup>26</sup> If  $P(e) = P(u) \times p_{ue}$ , the conditional probabilities  $P(ek)$  and  $P(e|\neg k)$  will be  $P(u|k) \times p_{ue}$  and  $P(u|\neg k) \times p_{ue}$ , respectively. As a result,

$$(i) \quad \Delta P_k^e = \frac{P(u|k) \times p_{ue} - P(u|\neg k) \times p_{ue}}{-P(u|\neg k)} = [P(u|k) - P(u|\neg k)] \times p_{ue}$$

We have deduced before that when  $k$  and  $u$  are the only potential causes of  $e$ , the general formula for causal power is the following:

$$(ii) \quad p_{ke} = \frac{\Delta P_k^e - [P(u|k) - P(u|\neg k)] \times p_{ue}}{1 - P(u|k) \times p_{ue}}$$

Substituting (i) for  $\Delta P_k^e$  in (ii) gives us the desired result:  $p_{ke} = 0$ .

But there is a price to be paid: we should either pose an **ambiguity**, or we **generalize** (but weaken) the analysis. On an ambiguity proposal, one could claim that although most generics of the form ‘*ks are e*’ are true, or acceptable, because of the causal power of *ks* to produce *e*, others are true, or acceptable, because of the causal power of *e*-ness to produce *k*. Because we believe that also (30) is true, or acceptable, in causal structure (ii), this won’t do however. Therefore, we think it is more appropriate to *generalize* the causal analysis.

Our proposal for a general analysis goes as follows (if we forget about impact):

- ‘*ks are e*’ is true, or acceptable, if and only if  $\Delta^*P_{k,(\neg u)}^e$  is high, due to a causal relation.<sup>27,28</sup>

But how does this more general analysis account for a generic of the form ‘*es are k*’, if *k* causes *e*, rather than the other way around? To answer that question, we will first define the probability that, given *e*, *e* is due to *k*,  $P(k \rightsquigarrow e|e)$ . After that we will show that under natural independence conditions this notion equals  $\Delta^*P_e^k$ .

Given that we derived before that in our causal structure  $k \rightarrow e \leftarrow u$ , objects of type *e* are caused by *k* with probability  $P(k) \times p_{ke}$ , the probability that, given *e*, *e* is due to *k* is

$$(32) \quad P(k \rightsquigarrow e|e) = \frac{P(k) \times p_{ke}}{P(e)} \quad 29$$

Notice that in causal structure  $k \rightarrow e \leftarrow u$  this value can be positive and high, while  $p_{ek} = 0$ . Although most generics of the form ‘Objects of type *k* are (generally) of type *e*’ are true because  $p_{ke}$  is high, others are true because  $P(e \rightsquigarrow k|k)$  is high. Observe that in contrast to  $p_{ke}$ , the value of  $P(e \rightsquigarrow k|k)$  depends crucially on the base rates of *k* and *e*, making the latter less ‘stable’ than the former.<sup>30</sup>

Next, we can show that in case one takes over Cheng’s independence assumptions by means of which she can estimate the causal power, one can show not only that

<sup>27</sup> Notice that we used  $\Delta^*P_{k,(\neg u)}^e$  instead of  $\Delta^*P_k^e$ . The reason is that we will assume that in case *k* is incompatible with the (disjunction of) alternative causes *u*, we will use  $\Delta^*P_{k,\neg u}^e$ , and  $\Delta^*P_k^e$  otherwise.

<sup>28</sup> To make things even less mysterious, it is perhaps wise to claim that the truth, or acceptability, of a generic sentence of the form ‘*ks are e*’ is always dependent on a specific causal background. For now we assume that context always makes clear what this specific causal background is.

<sup>29</sup> See Cheng et al. (2007). Notice that in case *k* is the only (potential) cause of *e*,  $p_{ke} = P(e|k)$ . In that case it immediately follows that  $P(k \rightsquigarrow e|e) = \frac{P(k) \times P(e|k)}{P(e)} = \frac{P(k \wedge e)}{P(e)} = P(k|e)$ .

<sup>30</sup> Perhaps this explains why generics expressed in the ‘causal order’ are more natural than the others.

$p_{ke} = \Delta^*P_k^e$ , but also that  $P(e \rightsquigarrow k|k) = \Delta^*P_k^e$ .<sup>31</sup> Because not only  $p_{ke}$ , but also  $P(e \rightsquigarrow k|k)$  holds for causal reasons, we have explained why *both* (29-a) and (29-b), represented by ‘*ks are e*’ and ‘*es are k*’, respectively, are true, or acceptable, if and only if  $\Delta^*P_k^e$  and  $\Delta^*P_e^k$ , respectively, are high due to a causal reason.

Suppose we have the following common cause structure:  $k \leftarrow u \rightarrow e$ . What about a generic of the form ‘*ks are e*’ like (30) ‘People with yellow fingers develop lung cancer’? How should we provide a causal analysis of this type of sentence in such a causal structure? It should be  $P(u \rightsquigarrow k|k) \times p_{ue}$ . Interestingly enough, in these circumstances this comes down to  $P(e|k)$ .<sup>32</sup>

Thus, given that  $P(u \rightsquigarrow k|k) \times p_{ue}$  measures the probability that *k* and *e* are produced by common cause *u*, the value of  $P(e|k)$  measures the same thing. As a result,  $\Delta^*P_k^e = \frac{P(e|k) - P(e|\neg k)}{1 - P(e|\neg k)}$  is a natural measure of correlation between *k* and *e* due to a causal reason. It follows that this case fits the general causal analysis of generic sentences.

<sup>31</sup> First we show that  $P(e|k) - P(e) = \frac{P(e)}{P(k)} \times [P(k|e) - P(k)]$ :

$$\begin{aligned} P(e|k) - P(e) &= \frac{P(k|e) \times P(e)}{P(k)} - \frac{P(k) \times P(e)}{P(k)} \\ &= \frac{1}{P(k)} \times [P(k|e) \times P(e)] - \frac{1}{P(k)} \times [P(k) \times P(e)], \\ &= \frac{1}{P(k)} \times [P(k|e) \times P(e) - P(k) \times P(e)] \\ &= \frac{1}{P(k)} \times P(e) \times [P(k|e) - P(k)]. \end{aligned}$$

Then one can show that  $\Delta^*P_k^e = \frac{P(e|k) - P(e|\neg k)}{1 - P(e|\neg k)} = \frac{P(e|k) - P(e)}{P(\neg e \wedge \neg k)}$ . Similarly,  $\Delta^*P_e^k = \frac{P(k|e) - P(k|\neg e)}{1 - P(k|\neg e)}$ . Given the above proof that  $P(e|k) - P(e) = \frac{P(e)}{P(k)} \times [P(k|e) - P(k)]$ , it follows that  $\Delta^*P_k^e = \frac{P(e)}{P(k)} \times \Delta^*P_e^k$ . But recall that under suitable independence conditions  $\Delta^*P_e^k = p_{ek}$ . It follows that  $\Delta^*P_k^e = \frac{P(e)}{P(k)} \times p_{ek} = \frac{P(e) \times p_{ek}}{P(k)} = P(e \rightsquigarrow k|k)$ .

<sup>32</sup> To see this, notice that

$$\begin{aligned} P(u \rightsquigarrow k|k) \times p_{ue} &= \frac{P(u) \times p_{uk}}{P(k)} \times P(e|u) \\ &\quad \text{because } e \text{ can only be caused by } u \\ &= \frac{P(u) \times P(k|u)}{P(i)} \times P(e|u) \\ &\quad \text{because } k \text{ can only be caused by } u \\ &= \frac{P(u \wedge k)}{P(k)} \times P(e|u) \\ &= P(u|k) \times P(e|u) \end{aligned}$$

Now we show that  $P(e|k) = P(u|k) \times P(e|u)$ . This we do as follows:

$$\begin{aligned} P(e \wedge k) &= \sum_{u \in U} P(k) \times P(u|k) \times P(e|u) \quad \text{by the chain rule} \\ P(e|k) \times P(k) &= P(k) \times \sum_{u \in U} P(u|k) \times P(e|u) \quad \text{exporting } P(k) \\ P(e|k) &= \sum_{u \in U} P(u|k) \times P(e|u) \quad \text{both divided by } P(k) \\ P(e|k) &= P(u|k) \times P(e|u) \quad \text{because only } U = 1 \text{ causes } e. \end{aligned}$$

## 6 Conclusion and Outlook

The goal of this paper was to see to what extent a causal power analysis of generics is defensible. We have seen that such an analysis is quite appealing in the following sense: it **explains** why under natural circumstances a generic of the form ‘*ks* are *e*’ is true iff the measure  $\Delta^*P_k^e$  is high, an analysis that was proposed before (by van Rooij and Schulz, in press) for empirical reasons. This explanation also has the conceptually appealing feature that it seems to align with our actual thinking. It forces us to look for suitable alternative potential causes and the relevant causal structures in which they are engaged. For instance, if two kinds both exhibit the same properties, it tries to come up with a common cause explanation. This forces one to look for ‘deeper’ analyses than a regularity analysis does. We feel, with Cartwright (1989), that this is also the way science works. Moreover, the causal analysis also gives rise to different empirical predictions in other than the ‘natural’ circumstances: under various conditions generics of the form ‘*ks* are *e*’ are seen to be true, or acceptable, although  $\Delta^*P_k^e$  is low. To account for the fact that for some examples where  $\Delta^*P_k^e$  is high, although  $p_{ke}$  is low, we have generalized the causal analysis. Moreover, we have seen that in various circumstances high causal power comes down to high (stable) conditional probability, which according to many authors (e.g. Cohen 1999) is the reason why most generics are true.<sup>33</sup>

In this paper we have been deliberately non-committing about whether our analysis of generics determines their truth conditions (if generics have them at all), or whether our analysis just involves their acceptability conditions. According to Haslanger (2010), Leslie (2013)—or so their proposals can be interpreted—a causal view should play a role only in *pragmatics*: the generic ‘Women are submissive’ should be avoided not so much because it is not true, but rather because it gives rise to the false suggestion that the generic is true for the wrong causal reasons, i.e., because of what it is to be a woman. One way to implement this suggestion is to claim that generics have truth-, or acceptability-, conditions based on correlations, but that many people assume that these correlations are the way they are because of their wrong *essentialist*’ reading of generics. We have suggested in Sect. 4.1 that if essences

play a key role in the causal interpretation of generics, causal power reduces naturally to conditional probability. Although this might lead to a somewhat stronger reading of generics than the one using  $\Delta^*P$ , it doesn’t lead to the much stronger interpretation that Haslanger and Leslie object to. Many proponents of a causal power view of regularities (e.g. Harré and Madden 1975; Ellis 1999), however, have something stronger in mind: the regularities are not just causal, but are taken to be (metaphysically) *necessary* (whatever that might mean exactly).<sup>34</sup> It is exactly against this latter strong — and we think wrong—essentialist view of generics that Haslanger (2010) and others warn us. Haslanger argues—just like Barth (1971) before her—that because generic sentences like ‘Women are submissive’ and ‘Bantus are lazy’ are taken to say something about the essence of, or of the real, women and Bantus, they have their malicious social impact: they introduce prejudices to children, strengthen existing ones, and are excellent strategic tools for propagandists because they are immune to counterexample: any non-submissive woman is not a real woman. We think, however, that once the connection between causal powers (or essentialism) and necessity is given up, some of Haslanger’s complaints against the use of generics lose their force. It still leaves open, however, the idea that causal powers should be used in pragmatics, to account for the appropriateness of generic sentences, rather than in semantics, to account for their truth (if generics have truth conditions at all).

## Compliance with Ethical Standards

**Conflicts of interest** Robert van Rooij declares that he has no conflict of interest. Katrin Schulz declares that she has no conflict of interest.

**Ethical Approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

<sup>33</sup> Another pleasing consequence is, that just like episodic sentences, generic sentences are on a causal power analysis predicted to be true just (if we take generics to have truth conditions) because a certain fact obtains. Because our analysis is compatible with the assumption that generics have truth-makers (the causal powers) that are independent of the base rates, we predict—in contrast to purely probabilistic analyses—that generics (can) express propositions and can be used in embedded contexts, like in ‘Countries that do not honor women’s rights, do not honor general human rights’.

<sup>34</sup> Anjum and Mumford (2010) advocate the position that a causal power approach is better not thought of in terms of necessity. They argue that the (neo)-Humean arguments against a causal power view evaporates when giving up this association. According to them, the weaker view—a view much in accordance with our approach—was defended as well by philosophers like Aquino and Geach.

## References

- Anjum RL, Mumford S (2010) A powerful theory of causation. In: Marmodoro A (ed) *The metaphysics of powers*. Routledge, London, pp 143–59
- Barth E (1971) *De Logica van de Lidwoorden in de Traditionele Filosofie*
- Bentham J (1824/1987) An introduction to the principles of morals and legislation. In: Mill JS, Bentham J (eds) *Utilitarianism and other essays*, Harmandsworth, Penguin
- Bird A (2007) *Nature's metaphysics laws and properties*. Oxford University Press, Oxford
- Carlson G (1977) Reference to kinds in English. Ph.D. dissertation, University of Massachusetts, Amherst
- Cartwright N (1989) *Nature's capacities and their measurement*. Oxford University Press, Oxford
- Chapman G, Robbins S (1990) Cue interaction in human contingency judgment. *Memory Cognit* 18:537–45
- Cheng PW (1997) From covariation to causation: a causal power theory. *Psychol Rev* 104:367–405
- Cheng PW (2000) Causality in the mind: estimating contextual and conjunctive causal power. In: Keil F, Wilson R (eds) *Explanation and cognition*. MIT Press, Cambridge, pp 227–253
- Cheng P, Novick L (2004) Assessing interactive causal influence. *Psychol Rev* 111:455–485
- Cheng PW, Novick LR, Liljeholm M, Ford C (2007) Explaining four psychological asymmetries in causal reasoning: implications of causal assumptions for coherence. In: Campbell JK et al (eds) *Topics in contemporary philosophy, causation and explanation*, vol 4. MIT Press, Cambridge, pp 1–32
- Chierchia G (1995) Individual level predicates as inherent generics. In: Carlson G, Pelletier F (eds) *The generic book*. Chicago University Press, Chicago, pp 176–223
- Cohen A (1999) *Think generic! The meaning and use of generic sentences*. CSLI Publications, Stanford
- Crupi V, Tentori K, Gonzalez M (2007) On bayesian measures of evidential support: theoretical and empirical issues. *Philos Sci* 74:229–52
- Danks D (manuscript) *The mathematics of causal capacities*. Carnegie Mellon University
- Danks D (2014) *Unifying the mind. Cognitive rerepresentations as graphical models*. MIT press, Cambridge
- Eckhardt R (1999) Normal objects, normal worlds, and the meaning of generic sentences. *J Semant* 16:237–278
- Ellis B (1999) Causal powers and laws of nature. In: Sankey H (ed) *Causation and laws of nature*. Kluwer Academic Publishers, Dordrecht, pp 19–34
- Fitelson B, Hitchcock C (2011) Probabilistic measures of causal strength. In: McKay Illari P, Russo F, Williamson J (eds) *Causality in the sciences*. Oxford University Press, New York, pp 600–627
- Gelman SA (2003) *The essential child*. Oxford University Press, New York
- Glymour C (2001) *The mind's arrows: bayes nets and graphical causal models in psychology*. MIT Press, Cambridge
- Good IJ (1961) A causal calculus I. *Br J Phil Sci* 11:305–318
- Goodman N (1954) *Fact, fiction, and forecast*. Harvard University Press, Cambridge
- Greenberg Y (2003) *Manifestations of genericity*. Routledge, London
- Halpern J (2016) *Actual causation*. MIT Press, Cambridge
- Harré R, Madden E (1975) *Causal powers: a theory of natural necessity*. Basic Blackwell, Oxford
- Haslanger S (2010) Ideology, generics, and common ground. In: Witt C (ed) *Feminist metaphysics: essays on the ontology of sex, gender and the self*. Springer, Dordrecht, pp 179–207
- Kahneman D, Wakker P, Sarin R (1997) Back to Bentham? Explorations of experienced utility. *Q J Econ* 112:375–405
- Keil FC (1989) *Concepts, kinds, and cognitive development*. Bradford Books/MIT Press, Cambridge
- Kenemy J, Oppenheim P (1952) Degrees of factual support. *Phil Sci* 19:307–324
- Krifka M, Francis J, Pelletier G, Carlson A, ter Meulen G, Chierchia G, Link G (1995) Genericity: an introduction. In: Carlson G, Jeffrey Pelletier F (eds) *The generic book*. University of Chicago Press, Chicago, pp 1–124
- Kripke S (1972/80) Naming and necessity. In: Davidson D, Harman G (eds), *Semantics of Natural Language*, Dordrecht, pp 253–355, 763–769
- Lawler JM (1973) *Studies in english generic*, university of michigan papers in linguistics 1: 1. University of Michigan Press, Ann Arbor
- Leslie SJ (2008) Generics: cognition and acquisition. *Philos Rev* 117:1–47
- Leslie SJ (2013) Essence and natural kinds: when science meets preschooler intuition. In: Gendler T, Hawthorne J (eds) *Oxford studies in epistemology*, vol 4. Oxford University Press, Oxford, pp 108–166
- Lewis D (1973a) Causation. *J Philos* 70:556–67
- Lewis D (1973b) *Counterfactuals*. Blackwell, Oxford
- Niiniluoto I, Tuomela R (1973) Theoretical concepts and hypothetico-inductive inference. Reidel, Dordrecht
- Pearl J (1988) *Probabilistic reasoning in intelligent systems*. Morgan Kaufman Publishers, Inc., San Mateo
- Pearl J (2000) *Causality: models, reasoning and inference*. Cambridge University Press, Cambridge
- Pearl J (2014) Interpretation and identification of causal mediation. *Psychol Methods* 19:459–481
- Preacher K, Kelley K (2011) Effect sizes measures for mediation models: quantitative strategies for communicating indirect effects. *Psychol Methods* 16:93–115
- Putnam H (1975) The meaning of meaning. In: Gunderson K (ed) *Language, mind and knowledge*. University of Minnesota Press, Minneapolis
- Quine WVO (1970) Natural Kinds. In: Rescher N et al (eds) *Essays in Honor of Carl G. Hempel*. D. Reidel, Dordrecht, pp 41–56
- Rehder B (2003) A causal-model theory of conceptual representation and categorization. *J Exp Psychol* 29:1141–1159
- Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: the effectiveness of reinforcement and non-reinforcement. In: Black A, Prokasy W (eds) *Classical conditioning II: current research and theory*. Appleton-Century-Crofts, New York, pp 64–69
- Ryle G (1949) *The concept of mind*. Hutchinson, Abingdon
- Schneider D (2004) *The psychology of stereotyping*. The Guilford Press, New York
- Shanks (1995) *The psychology of associative learning*. Cambridge University Press, Cambridge
- Shep MC (1958) Shall we count the living or the dead? *N Engl J Med* 259:1210–1214
- Shoemaker S (1980) Causality and properties. In: Van Inwagen P (ed) *Time and cause*, D. Reidel, Kufstein, pp 109–135
- Skyrms DB (1980) Causal necessity. A pragmatic investigation of the necessity of laws. Yale University Press, New Haven
- Sloman S (2005) *Causal models. How people think about the world and its alternatives*. Oxford University Press, Oxford
- Spirtes P, Glymour C, Scheines R (2000) *Causation, prediction, and search*, 2nd edn. MIT Press, Cambridge
- Strawson G (1989) *The secret connection: causation, realism and David Hume*. Oxford University Press, Oxford
- Tenenbaum J, Griffiths T (2001) The rational basis of representativeness. In: *Proceedings of the 23th annual conference of the cognitive science society*, pp 1036–1041

- Tentori K, Crupi V, Bonini N, Osherson D (2007) Comparison of confirmation measures. *Cognition* 103:1007–1019
- Tessler M, Goodman N (in press) The language of generalization. *Psychol Rev*
- Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science* 185:1124–1131
- van Rooij R (2017) Generics and typicality. In: *Proceedings of Sinn und Bedeutung 22*, Berlin
- van Rooij R, Schulz K (in press) Generics and typicality: a bounded rationality approach. *Linguist Philos*
- van Rooij R, Schulz K (to appear) Natural kinds and dispositions: a causal analysis. *Synthese*. <https://doi.org/10.1007/s11229-019-02184-y>
- Yuille A (2006) Augmented Rescorla-Wagner and maximum likelihood estimation. *Adv Neural Inf Process Syst* 15:1561–1568

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.