

# Probabilistic Theories of Causality: Singular or Plural?

Joonsung Kim

**Abstract**—Several prominent philosophers of science (Cartwright [2], Dupré [3], Hausman [8], [9], Hitchcock [10]) pose a problem for the probabilistic theory of causality called the unanimity theory. In this paper, I focus on Hitchcock's criticism of the unanimity theory. I critically examine Hitchcock's argument against the unanimity theory. After introducing the desirable features of the unanimity theory, I show that the unanimity theory too does the same job Hitchcock's theory is intended to do. I conclude that the criticism of the unanimity theory on the contrary reveals the versatility of the unanimity theory some philosophers of science and some computer scientists ([17], [21]) do not notice.

**Index Terms**—Binary, causality, contrast, disjunctive factors, ternary, population, probabilistic theory, unanimity.

## I. INTRODUCTION

Several leading philosophers of science (Cartwright [1]) have developed probabilistic theories of causality in 40 years. Suppes [21] and Good [7] championed probabilistic analyses of causality. Inspired by their analyses, Cartwright [1], Skyrms [18], Eells and Sober [4], and Sober [19] developed the probabilistic theory of causality. In particular, Eells [5] significantly articulated it, which is called the unanimity theory. Two features of the unanimity theory are worth noticing. First, the probabilistic theory of causality explicates a causal role of a factor  $X$  for another factor  $Y$  always relative to population  $P$  exemplifying a kind or type. The causal role of  $X$  for  $Y$  is different depending on what population (strictly speaking, population kind or type) we are considering (Eells [5]). For example, smoking may be a positive causal factor of lung cancer in population  $P$  exemplifying middle age, while smoking may be a neutral or negative causal factor of lung cancer in population  $P$  exemplifying all features of human beings. Second, the probabilistic theory of causality assesses causal significance of  $X$  for  $Y$  in each of the background contexts relative to population  $P$ .

A formally intriguing but uneasy criticism has been raised of the probabilistic theories of causality. This criticism, which might be seemingly devastating to the probabilistic theories of causality, is due to the problem of disjunctive factors. In order to assess the causal significance of a medicine for patients' recovery, a medical team divides the study group of patients into three treatment groups, and in turn provides the first group with placebo  $A$ , the second group with a moderate dose  $B$  and the third group with a strong dose  $C$ . Suppose that the probability of recovery  $Y$

given each of  $A$ ,  $B$ ,  $C$  is as follows:  $Pr(Y | A) = 0.2$ ,  $Pr(Y | B) = 0.4$ ,  $Pr(Y | C) = 0.9$ . The medical team wants to know whether the moderate dose  $B$  is a positive causal factor for the patients' recovery  $Y$ . This example is introduced by Hitchcock [8] who slightly modifies the original example Humphreys [10] first presented. According to the probabilistic theories of causality, the medical team compares the probability of  $Y$  in the presence of  $B$ , i.e.,  $Pr(Y | B)$ , with the probability of  $Y$  in the absence of  $B$ , i.e.,  $Pr(Y | \neg B)$  in which ' $\neg$ ' refers to 'negation'. Since  $\neg B$ , the absence of  $B$ , is equivalent to a disjunctive factor  $A \vee C$  (in which ' $\vee$ ' refers to 'or'), the medical team needs to assess the probability of  $Y$  in the presence of  $A \vee C$ , i.e.,  $Pr(Y | A \vee C)$ . In computing  $Pr(Y | A \vee C)$ , each disjunct,  $A$  and  $C$ , of the disjunctive factor  $A \vee C$  confers *different* probabilities on the factor  $Y$ . The problem then arises of how one identifies a single causally significant probability of the factor  $Y$  in the presence of the disjunctive factor  $A \vee C$ . Hitchcock [10] introduces two problems due to the problem of disjunctive factors, and argues that the probabilistic theories of causality cannot meet them. Humphreys [10], as far as I know, first posed the problem of disjunctive factors and the two problems due to it for the probabilistic theories of causality. Suppose that  $Pr(A)$  is equal to  $Pr(C)$ . Then,  $Pr(Y | \neg B) = [Pr(A) Pr(Y | A) + Pr(C) Pr(Y | C)] / [Pr(A) + Pr(C)] = [(0.5)(0.2) + (0.5)(0.9)] / [(0.5) + (0.5)] = 0.55$ . So  $Pr(Y | B) < Pr(Y | \neg B)$ , which tells us that  $B$  is a *negative* causal factor for  $Y$ . Hitchcock [10] claims that this casual claim conflicts with our intuition that the moderate dose has *positive* causal significance for patients' recovery. This is the first problem the probabilistic theories of causality confront. Again, suppose instead that  $Pr(A)$  is 0.6 and  $Pr(C)$  is 0.1. Then, contrary to the previous case,  $B$  is now a *positive* causal factor for  $Y$  since  $Pr(Y | B) > Pr(Y | \neg B)$ . Thus what causal significance  $B$  has for  $Y$  depends on the ratio of  $Pr(A)$  to  $Pr(C)$ . Hitchcock [10] finds it odd that the *objective* causal significance of  $B$  for  $Y$  depends on the ratio of  $Pr(A)$  to  $Pr(C)$ . This is the second problem he raises for the probabilistic theories of causality. Hitchcock [10] further argues that the unanimity theory, which is founded on the binary contrast, cannot convey information about complex relations of causal relevance such that, as doses of medicine change, so does the probability of patients' recovery change. Hitchcock claims that *ternary* contrast instead should be considered not only to meet the problem of disjunctive factors but also to explicate the complex relations of causal relevance. Inspired by Holland [9], Hitchcock claims that  $B$  is, with regard to  $Y$ , contrasted not with  $\neg B$  but with a specific alternative to  $B$ , which is the ternary contrast. He compares  $Pr(Y | B)$  and  $Pr(Y | A)$ , and  $Pr(Y | B)$  and  $Pr(Y | C)$ , which are the ternary relations.  $B$  is a positive causal factor for  $Y$  relative to  $A$ ,  $Pr(Y | B) > Pr(Y |$

Manuscript received April 8, 2012; revised May 1, 2012. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2011-327-A00140).

Joonsung Kim is with the Department of Philosophy, Myongji University, Seoul, Korea (ROK) (e-mail: jkim30@mju.ac.kr).

A), whereas  $B$  is a negative causal factor for  $Y$  relative to  $C$ ,  $Pr(Y | B) < Pr(Y | A)$ . We also see that as doses of medicine change, so does the probability of patients' recovery change:  $Pr(Y | A) < Pr(Y | B) < Pr(Y | C)$ . Hitchcock formally generalizes these complex relations of causal relevance in terms of a conditional probability distribution function  $f(x) = Pr(Y | X = x)$  where  $X$  is a random variable standing for doses of medicine.  $Y$  is of course also a random variable. In this example, Hitchcock considers only two cases of effect, i.e., recovery and non-recovery, so that  $Y$  has only two values 1 and 0. The ternary theory introduces a conditional probability distribution function  $f_i(x) = Pr(Y = y | X = x \& K_i)$ , and contrasts different values of the random variable  $X$  with regard to values of a random variable  $Y$ . " $i$ " of  $f_i(x)$  represents each  $i$  of the background contexts  $K_i$ , so that the function  $f_i(x)$  may have different shapes, depending on what background context it is relative to. The ternary theory (Hitchcock [10]) is a generalization of Holland's [11] interpretation of causal relevance such that  $X$  is a positive, negative, or neutral cause of  $Y$  with respect to an alternative to  $X$ . Suppose that the values of the random variable  $X$  are doses of medicine determined as the result of a random experiment, values of a random variable  $Y$  are recovery or non-recovery, and the random experiment is relative to, for example, a background context  $K_2$ . If the probability of  $Y$  given  $X = 2$  (e.g., a moderate dose of medicine) is greater than the probability of  $Y$  given  $X = 1$  (e.g., a placebo), then  $X = 2$  tends to cause  $Y$  when compared with  $X = 1$ . If the probability of  $Y$  given  $X = 3$  (e.g., a strong dose of medicine) is greater than the probability of  $Y$  given  $X = 2$ , then  $X = 3$  tends to cause  $Y$  when compared with  $X = 2$ . Thus the ternary theory allegedly meets the two problems by showing that the moderate dose of medicine is a positive causal factor for the patients' recovery without depending on  $Pr(X_1)$  and  $Pr(X_3)$ . Hitchcock goes further. The function  $f_i(x)$  has a shape of probability increasing relative to the background context  $K_2$  such that  $f_2(1) < f_2(2)$  and  $f_2(2) < f_2(3)$ , assuming that  $f_2(1) = Pr(Y | X = 1 \& K_2) = 0.2$ ,  $f_2(2) = Pr(Y | X = 2 \& K_2) = 0.4$  and  $f_2(3) = Pr(Y | X = 3 \& K_2) = 0.9$ .  $y$  of  $Y=y$ , which represents different non-negative values of the random variable  $Y$ , will not appear in  $f_2(x)$ . For, in this example, Hitchcock considers only two cases of effect, i.e., recovery and non-recovery, so that  $Y$  has only two values 1 and 0. If the above random experiment is relative to another background context, then  $f_i(x)$  may have a shape of probability decreasing or a shape of probability not changing. The relations,  $f_2(1) < f_2(2)$  and  $f_2(2) < f_2(3)$ , convey the information about the function  $f_i(x)$  such that the probability of  $Y$  increases from  $X = 1$  through  $X = 2$  to  $X = 3$ . Hitchcock [8] claims that only the ternary theory conveys the information about the function  $f_i(x)$ , and is superior to the unanimity theory.

In this paper I shall first introduce the probabilistic theory of causality which is called the unanimity theory. Second, I shall show how the unanimity theory too conveys the information about the complex relations of causal relevance the ternary theory is intended to do. Third, I shall argue that the unanimity theory and the ternary theory both carve up the same causal structure in two formally different but conceptually consistent ways, while revealing the desirable

feature of the unanimity theory.

## II. THE PROBABILISTIC THEORY OF CAUSALITY: THE UNANIMITY THEORY<sup>1</sup>

A type-level relation between a factor  $X$  and another factor  $Y$  expresses causal tendency of a factor  $X$  for another factor  $Y$ . A question arises about what it means to say that  $X$  has causal tendency for  $Y$ . The basic idea of theories of probabilistic causation is that causal tendency is explicable in terms of probabilistic relevance, that is, the comparison between the probability of  $Y$  in the presence of  $X$  and the probability of  $Y$  in the absence of  $X$ . The probabilistic relevance indicates that causal tendency of  $X$  for  $Y$  should be relative to "some individual, individuals, or population" (Eells [5]). For, whatever an interpretation of probability may be, we cannot consider probability values without applying the (conditional) probability to a population. The relativity of probability values to populations allows us to consider the relativity of probabilistic relevance to populations, and of causal tendency to populations. The basic idea of type-level theories of probabilistic causation is "A property-level probabilistic causal claim must be made relative to a particular, token, population" (Eells [5]).

What does it mean to say that causal tendency of  $X$  for  $Y$  is relative to a population? It means that causal tendency of  $X$  for  $Y$  "can vary from population to population and from individual to individual" (Eells [5]). In order to understand this answer, we need to notice that the token population must be considered as the relevant token population of "a certain given kind": "A token population is always a token of many types, or kinds, and the kind of causal significance a factor  $X$  for a factor  $Y$ , in a population  $P$ , can depend on what kind,  $Q$ , of population we think of  $P$  as exemplifying" (Eells [5]). This point comes from the fact that "probability values can depend on what kind we associate with a given population"; more strictly speaking, the fact that probability values are explicated in terms of "frequencies in hypothetical populations of the relevant kind" (Eells [5]).

Consider an actual population  $P$ . We can consider the population  $P$  as the result of an experiment that could be repeated. Suppose that we are concerned with  $Pr(Y | X)$  and  $Pr(Y | -X)$  in the experiment. The experiment can proceed as follows. Let us provide a set of individuals in the population  $P$  with a certain "distribution of initial conditions" (Eells [5]). The initial conditions are "factors that are causally relevant to the way individuals in  $P$  are, with respect to having or lacking" the factors  $X$  and  $Y$ . The factors  $X$  and  $Y$  are not the members of these initial conditions. If there are  $k$  initial conditions, e.g.  $I_1, \dots, I_k$ , then a distribution of initial conditions has the frequencies of the  $2^k$  possible combinations of  $I_k$ 's and  $-I_k$ 's. A distribution of initial conditions is called "an experimental set-up," which gives the individuals in the population  $P$  "propensities of various strengths to have or lack" the factors  $X$  and  $Y$ . If we repeat the experiment under the experimental set-up, then we "distribute these conditions in exactly the same way, in a (possibly) different population of individuals" (Eells [5]). So

<sup>1</sup> This section is excerpted from Kim [14].

the result of the experiment is the resulting conditional frequencies that involve  $X$  and  $Y$  in  $P$ . The resulting conditional frequencies are supposed to be different from one experiment, or hypothetical population, to another experiment, or another hypothetical population, assuming that indeterminism holds in the world. That is, these propensities will be different “from individual to individual, depending on which of the initial conditions are present and absent in the individuals” (Eells [5]). The distribution of initial conditions is a *kind* or *type*  $Q$  population  $P$  exemplifies.

What is important is that the probability of  $Y$  in the presence of  $X$  and the probability of  $Y$  in the absence of  $X$  are different, depending on which kind or type  $Q$  a population  $P$  is considered as exemplifying, in other words, under which set of initial conditions, or which experimental set-up an experiment is repeated. Consider, for example, the probability of a coin coming up with a head when it is tossed. Then, the coin tossing is an experiment that is endowed with a distribution of initial conditions, e.g.,  $I_1, \dots, I_k$ . Let 20 trials on which a coin is tossed be an actual population  $P$ . We can envisage the sequence of hypothetical populations,  $P_1(=P), P_2, \dots, P_n$ , that would result from conducting the experiment finitely many times. Then, the resulting conditional frequencies of a coin coming up with a head when it is tossed are expected to be different from a hypothetical population (e.g.,  $P_1$ ) to another hypothetical population (e.g.,  $P_2$ ). Note that the probability of a coin coming up with a head on a coin tossing is identified with the average hypothetical relative frequency of a coin coming up with a head on a coin tossing across infinitely many different experiments (infinitely many finite frequencies  $P_1(=P), P_2, \dots$ ) of the random experiment setting  $Q$ , that is, across infinitely many different experiments endowed with a set of initial conditions, i.e., a kind or type  $Q$ . (The average hypothetical relative frequency is intended to indicate a modality of causal tendency by washing out the effect of accidental coincidences in an actual and single population.) What is important is that the token populations  $P_1(=P), P_2, \dots, P_n$  are considered as exemplifying a kind or type  $Q$ , i.e., a distribution of the initial conditions. On the other hand, consider a different set of initial conditions, e.g.,  $J_1, \dots, J_m$ , and associate it with the above infinitely many different experiments,  $P_1(=P), P_2, \dots$ . The different set of initial conditions is a different kind or type  $Q^*$  the token populations,  $P_1(=P), P_2, \dots, P_n$ , exemplify. That is, the infinitely many finite frequencies,  $P_1(=P), P_2, \dots$ , have both  $Q$  and  $Q^*$ . Then, the probability of a coin coming up with a head on a coin tossing is expected to be a *different* average hypothetical relative frequency of a coin coming up with a head on a coin tossing across infinitely many different experiments (infinitely many finite frequencies  $P_1(=P), P_2, \dots$ ) of the different random experiment setting  $Q^*$ . Thus, the conditional probability values, which are interpreted as hypothetical relative frequencies, depend on which set of *initial conditions*, or which *population type* a population is considered as exemplifying.

The relativity of the (conditional) probability values to a population exemplifying a kind or type allows us to consider the relativity of probabilistic relevance to a population (i.e., the probability of  $Y$  in the presence of  $X$  and the probability of

$Y$  in the absence of  $X$  relative to a population  $P$  exemplifying a type or kind  $Q$ ), which is intended to explicate *causal tendency* of  $X$  for  $Y$ . Consider an example of causal tendency of smoking for lung cancer relative to a population  $P$ . Then, we can consider two different kinds or types  $Q$  and  $Q^*$ , i.e., two different sets of initial conditions, or two different experimental settings. Consider, for example (Eells [5]), the actual human population that exemplifies both the kind *human*  $Q$  and the more complicated kind  $Q^*$  described as everyone’s actual causal conditions. Then, relative to the token population of  $Q$ , smoking is a positive causal factor for lung cancer in the token population. But relative to the token population of  $Q^*$ , smoking is a negative or neutral causal factor for lung cancer in the token population. Thus, causal tendency of smoking for lung cancer may be different, depending on which of the two different kinds or types  $Q$  and  $Q^*$  the token population  $P$  is considered as exemplifying.

Let me summarize the heart of the theory of type-level probabilistic causation. Type-level probabilistic causation is a relation among four components: a causal factor  $X$ , an effect factor  $Y$ , a token population  $P$  within which the first factor is some kind of cause of the second factor, and a kind or type  $Q$  that is associated with the given token population  $P$  (Eells [5]).

The relativity of probability values to a population allows us to consider the relativity of probabilistic relevance to a population  $P$  of a type  $Q$ . A question arises about how probabilistic relevance explicates causal tendency of  $X$  for  $Y$ . Consider a causal claim, “Smoking,  $X$ , causes lung cancer,  $Y$ , relative to a population  $P$  of  $Q$ .” We can envisage that there are, in a population, numerous factors including smoking,  $X$ , (e.g., a genetic factor, polluted air and so on) causally relevant to lung cancer,  $Y$ . Then, the above question is paraphrased as follows: How can we assess causal tendency (or causal significance) of  $X$  for  $Y$  in a population in terms of probabilistic relevance?

The theory of probabilistic causation answers this question as follows: In assessing causal significance of  $X$  for  $Y$ , we should hold fixed, in background contexts, all the factors, independently of  $X$  (in the sense that  $X$  is not a cause of those factors), causally relevant to  $Y$ . Consider, for example, residents in the city whose air is polluted (Skyrms [18], Eells [5]). In what follows, the concept of causal independence will be used in the same sense. They will refrain from smoking, since they are afraid of being exposed to both of the two hazard factors, that is, smoking and the polluted air. On the other hand, residents in the country do not hesitate to smoke, since they are not worried about the double hazard factors. So living in the country  $Z$  brings about both smoking  $X$  and being healthy  $-Y$  ( $Y$ : getting lung cancer). Then, even though smoking  $X$  does not cause the country residents’ health  $-Y$  (i.e., they do not get lung cancer), it is positively correlated with the country residents’ health. Smoking  $X$  lowers the probability of lung cancer  $Y$ , since smoking  $X$  is positively correlated with living in the country  $Z$ , which is negatively correlated with lung cancer  $Y$ . So it seems that the country residents’ smoking is beneficial to their health. But notice that living in the country  $Z$  plays a common causal role in bringing about the country residents’ smoking  $X$  and the country residents’ health  $-Y$ . So the relation between smoking

and health is a *spurious correlation*. Let us hold fixed living in the country  $Z$  (which is, independently of smoking  $X$ , causally relevant to getting lung cancer  $Y$ ) in background contexts:  $Pr(Y | X \& Z) > Pr(Y | -X \& Z)$  and  $Pr(Y | X \& -Z) > Pr(Y | -X \& -Z)$ . Then, positive probabilistic relevance of  $X$  to  $-Y$  disappears in the subpopulation  $Z$ , and  $X$  is positively probabilistically relevant to  $Y$ .

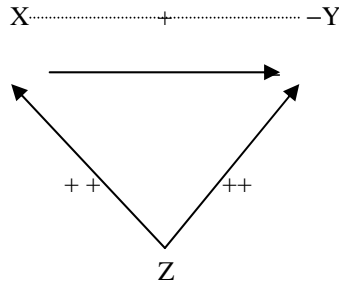


Fig.1. Common cause and spurious correlation.

As the Fig. 1 indicates, spurious correlation arises when a third factor is correlated with a causal factor and an effect factor. By holding fixed the third factor (which is, independently of  $X$ , causally relevant to  $Y$ ) in a background context,  $Pr(Y | X \& Z) > Pr(Y | -X \& Z)$  and  $Pr(Y | X \& -Z) > Pr(Y | -X \& -Z)$ , and we can assess true causal significance of  $X$  for  $Y$ . This is a crucial reason why all the factors, which are, independently of a factor  $X$ , causally relevant to another factor  $Y$ , should be held fixed in background contexts.

Let me introduce a formal generalization of background contexts. (Eells [5]) Let all the factors, independently of  $X$ , causally relevant to  $Y$ , be  $F_1, \dots, F_n$ . The factors  $F_1, \dots, F_n$  are supposed to be held fixed positively or negatively as follows. Since there are  $n$  of these factors, there are  $2^n$  ways in which they can be held fixed,  $2^n$  maximal conjunctions. Of these  $2^n$  maximal conjunctions, let  $K_1, \dots, K_m$  be conjunctions that have non-zero probability both in conjunction with  $X$  and in conjunction with  $-X$ : for  $i = 1, \dots, m$ ,  $Pr(X \& K_i) > 0$  and  $Pr(-X \& K_i) > 0$ . These  $K_i$ 's are called "background contexts." (*Ibid.*) Thus, the theory of probabilistic causation requires that all the factors, causally independent of  $X$ , causally relevant to  $Y$  should be held fixed in background contexts  $K_i$ . The unanimity theory explicates causal tendency of  $X$  for  $Y$  relative to a population  $P$  of a population type  $Q$  as follows:  $X$  is a positive, neutral or negative causal factor for  $Y$  in a population  $P$  of  $Q$  if and only if  $Pr(Y | X \& K_i) >, =, < Pr(Y | -X \& K_i)$  for each  $i$ .

This inequality or equality must hold for *each* of the background contexts  $K_i$ , which is called the condition of "contextual unanimity." Otherwise,  $X$  is *mixed* for  $Y$ . That is, a factor  $X$  is a mixed causal factor for a factor  $Y$  if  $X$  does not unanimously have positive, negative or, neutral probabilistic significance for  $Y$  when all the factors, independent of  $X$ , causally relevant to  $Y$  are held fixed in background contexts  $K_i$ . Thus, the unanimity theory explicates *four* types of causal relevance (i.e., positive, negative, neutral and mixed causal relevance) in terms of probabilistic relevance across the background contexts. The case of mixed causal relevance

needs more explanation. Suppose, for example, that ingesting acid (or alkali) raises the probability of death. Suppose that among the individuals who ingest acid (or alkali), there are individuals who have already ingested alkali (or acid). Then, for them, ingesting acid (or alkali) does not increase the probability of death. Cartwright [1] calls this "causal interaction." The case raises a question about what kind of causal significance theories of probabilistic causation assign to ingesting acid (or alkali) for death. The unanimity theory answers this question as follows. In the general population, i.e., in the whole population including those who have already ingested alkali (or acid) and those who have not, ingesting acid (or alkali) is *mixed* for death, while, in the subpopulation of individuals who have already ingested alkali (or acid), ingesting acid (or alkali) is a *negative* causal factor for death. Thus, the unanimity theory allows us to consider causal truth in subpopulations, and allows us to see "reversals of causal roles across subpopulations" in the above case (Eells [4]). If a theory of probabilistic causation claims that ingesting acid (or alkali) is a positive causal factor for death in the general population, then it will neglect crucial aspects of causal truth. Consider, for example, that agents are willing to make a decision. Then, the agents are in general concerned with in which of the subpopulations they are located. If so, then causal truth not in the general population but in subpopulations should guide them to make a rational decision. The unanimity theory provides the agents with the crucial aspects of causal truth in subpopulations, and can guide their rational decision.

Notice how the unanimity theory explicates the reversal of causal role across subpopulations in the above case of ingesting acid (or alkali). The unanimity theory requires that we should hold fixed ingesting alkali (or acid) in the general population that combines those who have already ingested alkali (or acid) and those who have not. Then, we can conclude that ingesting acid (or alkali) is mixed for death in the general population. The factor of ingesting alkali (or acid) is called an "interactive factor." The interactive factor is the factor with which a factor  $X$  *interacts* in the production of another factor  $Y$ . So the unanimity theory requires us to *hold fixed* not just positive, negative and mixed causal factors for  $Y$  but also the *factors* with which  $X$  *interacts* in the production of  $Y$ . By holding fixed the *interactive* factors (e.g., ingesting alkali (or acid)), we can assess causal significance of the *uncombined* causal factor (e.g., ingesting acid (or alkali)) for death: Ingesting acid (or alkali) is mixed for death in the general population, while it is a negative causal factor for death in the subpopulation of those who have already ingested alkali (or acid). There is a special case of causal interaction in which combined causal factors poses a problem for the unanimity theory. Notice that when  $X$  interacts with  $F$  with regard to  $Y$ , the interactive factor  $F$  is not always a positive, negative, or mixed causal factor for  $Y$ . There is the case in which an interactive factor  $F$  is causally *neutral* for  $Y$ . Consider, for example, that a physiological factor  $F$  is a *neutral* factor for lung cancer  $Y$ . And the physiological factor  $F$  is also positively causally relevant to smoking  $X$ . Smoking  $X$  is mixed for lung cancer  $Y$  since  $X$  interacts with  $F$  with regard to  $Y$ . (See below Fig.1)

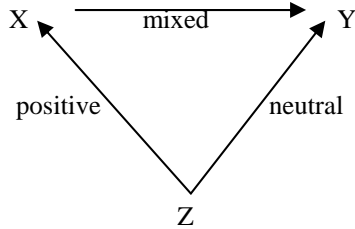


Fig. 2. Mixed causal relevance.

As we will see in a moment, an intermediate factor smoking  $X$  should not be held fixed. Since  $F$  is causally neutral for  $Y$ , it *prima facie* seems that, in assessing the causal significance of  $X$  for  $Y$ ,  $F$  does not definitely need to be held fixed. But if  $F$  is not held fixed, then this case can conceal that  $X$  is mixed for  $Y$  (since  $X$  interacts with  $F$  with regard to  $Y$ ) in assessing the causal significance of  $X$  for  $Y$ . So even though  $F$  is causally neutral for  $Y$ , it should be held fixed since  $X$  interacts with it with regard to  $Y$  (See Eells [5], for details). This is the crucial kind of case in which interactive factors should be held fixed.

Let me summarize the unanimity theory: Relative to a population  $P$  of a population type  $Q$ , all the factors, causally independent of  $X$  (i.e.,  $X$  is not a cause of those factors), causally relevant to  $Y$  (i.e., all the factors which are positive, negative, neutral, or mixed causal factors for  $Y$ ) should be held fixed in background contexts  $K_i$ . Also, all the factors that are causally independent of  $X$  and with which  $X$  causally interacts with regard to  $Y$  should be held fixed in background contexts  $K_i$ . Then,  $X$  is a positive, neutral or negative causal factor for  $Y$  in population  $P$  of  $Q$  if and only if  $Pr(Y | X \& K_i) >, =, < Pr(Y | -X \& K_i)$  for each  $i$ . Otherwise,  $X$  is *mixed* for  $Y$ .

At this point, it is worth noticing that the theory of probabilistic causation confronts a problem of circularity. In order to judge that one factor is positively causally relevant to another factor, we need to hold fixed all the factors, independently of the first factor, *causally relevant* to the second factor in background contexts. Thus, the *causal relevance* of the first factor to the second factor is explicated in terms of the *causal relevance* of the third factor to the second factor. So the theory of probabilistic causation confronts circularity. But two points are worth noticing. First, the theory of probabilistic causation is a theory of the “relation between probability and causation” (Cartwright [1], [2], Eells [5]). In other words, it is not a theory that *reduces* causation to probability but a theory about relations between probabilistic causal relations. Of course, a question arises about why causal relevance is associated with probabilistic relations. We already discussed an answer to this question in the previous section. Cartwright proposes a profound ontological answer to the question. Capacities associate causal relevance with probabilistic relations. I will discuss Cartwright’s accounts of capacities in chapter 2. Second, Skyrms [18] makes the following point. “ $X$  is a causal factor for  $Y$ ” is characterized not in terms of  $X$  being causally relevant to  $Y$  but in terms of the other factors than  $X$ . So the circularity with the theory of probabilistic causation is not so vicious as it could be.

### III. SEEMINGLY MULTIPLE WAYS BUT ONE WAY OF CONTRAST

The unanimity theory is also a ternary theory: it says that a factor  $X$  is a causal factor for another factor  $Y$  *relative to a population  $P$*  (Eells and Sober [6]). Strictly speaking, the unanimity theory is a quandary theory which says that a factor  $X$  is a causal factor for another factor  $Y$  relative to a population  $P$  exemplifying a kind, or type  $Q$ . Hitchcock’s ternary theory too is the same as the unanimity theory in that it too carves up causal structure relative to a population  $P$  exemplifying a kind, or type  $Q$ . Hitchcock did not notice how versatile the relativity of causal roles to population would be while I do. The causal significance of  $X$  for  $Y$  depends on which population we are considering. This is understood in two ways. First, a population  $P$  always exemplifies a population type  $Q$ . The causal significance of  $X$  for  $Y$  depends on which population type  $Q$  the population  $P$  is taken to exemplify. For example, smoking may have a positive causal significance for lung cancer in a population of middle-aged human beings. But smoking may not have positive causal significance for lung cancer in a population of teen-aged human beings. Second, a population  $P$ , in which a factor is a causal factor for another factor, is basically taken as a homogeneous subpopulation. Causal role may be different, depending on which subpopulation we are considering. For example, if  $X$  is a positive causal factor for  $Y$  in a homogeneous subpopulation, then  $X$  may be a negative causal factor for  $Y$  in another homogeneous subpopulation. If this causal information is true, then  $X$  is causally mixed for  $Y$  in a subpopulation into which the two subpopulations are combined. Let us see how this feature of the probabilistic theory of causality conveys information about the function  $f_i(x)$  the ternary theory alone allegedly does.

Consider, for brevity, only the three cases of the experiment relative to a background context  $K_2$  in terms of Hitchcock’s ternary theory introduced in the previous section,  $f_2(1) = Pr(Y | X = 1 \& K_2) = 0.2, f_2(2) = Pr(Y | X = 2 \& K_2) = 0.4$  and  $f_2(3) = Pr(Y | X = 3 \& K_2) = 0.9$ . Let  $X = 1, X = 2, X = 3$  be in turn  $X_1, X_2, X_3$ , which constitute a partition of doses of medicine. See Fig. 3.

	-Y	-Y	-Y	
0.4				0.9
0.2	Y	Y	Y	
	$X_1$	$X_2$	$X_3$	

Fig. 3. A partition of doses of medicine.

According to the unanimity theory, the relations between  $X_1, X_2, X_3$  and  $Y$  relative to the background context  $K_2$  are  $Pr(Y | X_1 \& K_2), Pr(Y | X_2 \& K_2), Pr(Y | X_3 \& K_2)$ . These three conditional probabilities are in turn equivalent to  $f_2(1) =$

$Pr(Y | X=1 \& K_2) = 0.2$ ,  $f_2(2) = Pr(Y | X=2 \& K_2) = 0.4$ ,  $f_2(3) = Pr(Y | X=3 \& K_2) = 0.9$ . Consider, relative to  $K_2$ , a subpopulation whose individuals have the property of  $X_1$  or  $X_2$ . That is, a population type of the subpopulation is  $X_1$  or  $X_2$ . (The shaded parts in Figure 2 are the relations between  $X_1$ ,  $X_2$  and  $Y$  in the subpopulation  $X_1$  or  $X_2$ .)

Let the subpopulation relative to  $K_2$  be  $K_2^1$ . In the subpopulation  $K_2^1$ ,  $X_2$  is not only the absence of  $X_1$  but also only alternative to  $X_1$ . Therefore, in the subpopulation  $K_2^1$ ,  $Pr(Y | X_1) < Pr(Y | -X_1)$ , that is,  $Pr(Y | X_1) < Pr(Y | X_2)$ . Again, consider a subpopulation whose individuals have the property of  $X_2$  or  $X_3$ . (The shaded parts in Figure 3 are the relations between  $X_1$ ,  $X_2$ , and  $Y$  in the subpopulation  $X_2$  or  $X_3$ .)

Let the subpopulation be  $K_2^2$ . In the subpopulation  $K_2^2$ ,  $X_3$  is not just the absence of  $X_2$  but also only alternative to  $X_2$ . Therefore, in the subpopulation  $K_2^2$ ,  $Pr(Y | X_2) < Pr(Y | -X_2)$ , that is,  $Pr(Y | X_2) < Pr(Y | X_3)$ . Again, consider a subpopulation whose individuals have the property of  $X_1$  or  $X_3$ . (The shaded parts in Figure 4 are the relations between  $X_1$ ,  $X_3$ , and  $Y$  in the subpopulation  $X_1$  or  $X_3$ .)

Let the subpopulation be  $K_2^3$ . In the subpopulation  $K_2^3$ ,  $X_3$  is not just the absence of  $X_1$  but also only alternative to  $X_1$ . In the subpopulation  $K_2^3$ ,  $Pr(Y | X_1) < Pr(Y | -X_1)$ , that is,  $Pr(Y | X_1) < Pr(Y | X_3)$ . Let us follow the transition of the probability of  $Y$  from the subpopulation  $K_2^1$  through the subpopulation  $K_2^2$  to the subpopulation  $K_2^3$ . Then it is easy to see that the probability of  $Y$  increases from  $X_1$  through  $X_2$  to  $X_3$ . This conveys the information exactly about  $f_2(x)$  such that  $f_2(1) < f_2(2)$  and  $f_2(2) < f_2(3)$ . By considering *each of the three subpopulations*  $K_2^1$ ,  $K_2^2$ ,  $K_2^3$  as the third relatum, the unanimity theory shows that the comparison between  $Pr(Y | X)$  and  $Pr(Y | -X)$  conveys the information about  $f_2(x)$ .

#### IV. A TOPOGRAPHY OF PROBABILISTIC THEORIES OF CAUSALITY: THE VERSATILITY OF THE UNANIMITY THEORY

Hitchcock does not notice that the unanimity theory is developed in the context of a *three-place* theory of causality: it says that a factor is a causal factor for another factor *relative to a population of a certain type or kind*. Therefore, the theory is, as it should be, sensitive to the mechanism by which subjects are assigned to the three treatment groups. This sensitivity is not an objection to the theory but rather clearly reveals a desirable feature of the unanimity theory. Several philosophers (Hausman [8], [9], Schaffer [17]), which have discussed what should be contrasted as causal relata, argue that ternary (or quandary) contrast meets several crucial problems with theories of causation allegedly due to binary contrast. They at least implicitly credit their arguments for quandary contrast to Hitchcock's [10] ternary probabilistic theory of causality. My discussions in this paper may lead them to reconsider the relation between binary contrast and ternary contrast. At least as far as causal relevance is relative to population, the ternary contrast is founded on the binary contrast. Hitchcock does not notice that Eells' *solution* is developed in the context of a *three-place* theory of causality: it says that a factor is a causal factor for another factor *relative to a population of a certain type or kind*. Therefore, the theory is, as it should be,

sensitive to the mechanism by which subjects are assigned to the three treatment groups. This sensitivity is not an objection to the theory but rather clearly reveals a desirable feature of the unanimity theory. In the footnote of his paper, Hitchcock mentions Eells' response to his paper, and notes that the unanimity theory is three-place theory. But it is interesting that Hitchcock disregards the desirable feature of the unanimity theory. Let us generalize the relation between the ternary theory and the unanimity theory. The unanimity theory and the ternary theory both explicate the same relations of causal relevance in each  $i$  of  $K_i$  in two different ways as follows. In each single background context, the ternary theory considers, as a third relatum, a third non-negative value that a random variable takes, while, in each single background context, the unanimity theory considers subpopulation as a third relatum. There is no conflict between the two ways. In each single  $i$  of  $K_i$ , the unanimity theory and the ternary theory both carve up the same causal structure in two conceptually different but consistent ways. It is worth noticing that Suppes [21] makes the same point. Suppes presents a quantitative probabilistic theory of causality as well as the probabilistic theory of causality. The quantitative theory is conceptually equivalent to the probabilistic theory of causality but different from the probabilistic theory of causality only in the form of analysis. The quantitative theory coincides with the ternary theory in the sense that probabilities are defined over random variables rather than events. The conceptually equivalent relation between the two theories too confirms that the unanimity theory coheres with the ternary theory.

Now I shall deliver a strong message to those who disregard the unanimity theory in favor of Hitchcock's ternary theory. Contrary to Hitchcock's criticism of the unanimity theory, the unanimity theory is more versatile than the ternary theory. The ternary contrast compares a factor  $X$  and a single third factor  $W$  that is a member of the set of the other factors than  $X$ . But the ternary contrast cannot compare  $X$  and  $-X$ , even though it is intended to compare them. Since  $-X$  indicates a complement of a singleton set including only  $X$ ,  $-X$  refers to any of the other factors than  $X$  (i.e.,  $W_1, W_2, W_3, \dots, W_k$ ), which can be *disjunctively* conjoined in different ways (i.e.,  $W_1 \vee W_2, W_1 \vee W_3, W_1 \vee W_2 \vee W_3, \dots$ ). This is the very problem of disjunctive factors, which is the one we should have met. Hitchcock's theory is motivated by the problems with probabilistic theories of causality *due to* the problem of disjunctive factors. But Hitchcock's theory cannot solve the problem of disjunctive factors *per se*. In the experimental situation of dose of medicine, the ternary theory clearly cannot compare  $B$  and  $A \vee C$  with regard to patients' recovery  $Y$  since his theory does not consider a disjunctive factor. His theory never assesses causal significance of a disjunctive factor for another factor. Hitchcock would say that he *dissolves* the problem of disjunctive factors. Then his response is merely a rhetorical defense. Hitchcock merely bypasses the problem of disjunctive factors since the problem of disjunctive factors *per se* still remains intact, as far as we keep considering it as a genuine problem of the probabilistic theories of causality.

Against my criticism of the ternary, Hitchcock would

claim that the ternary theory is more versatile than the unanimity theory. Hitchcock [10] claims that the ternary theory is more practical than the unanimity theory since it provides information about causal significance of moderate dose for patients. The information helps doctors to make a rational decision. But the unanimity theory also provides the very information for doctors to do so. Hitchcock should notice that the information about causal significance of the disjunctive factor  $A \vee C$  for patients' recovery too guides doctors to make a rational decision. As far as Hitchcock bypasses answering the problem of disjunctive factors *per se*, he should instead accept Eells' solution if he agrees with my argument for Eells' solution founded on the unanimity theory.

Hitchcock could claim that in the above experiment situation causal significance of the disjunctive factor  $A \vee C$  for  $Y$  impairs our intuition that moderate dose of medicine is the best treatment to patients. So it could be a strategy in that experimental situation to bypass assessing causal significance of  $A \vee C$  for  $Y$ . But the situation is a particular one that meets the alleged ordinary intuition. We need to see that our ordinary intuition is different in the other situations. We should not always consider the moderate dose of medicine as the best treatment. Considering possibly many different situations, it is truly necessary to assess causal significance of disjunctive factors for another factor. Hitchcock also claimed that it is not admissible that objective causal significance of a factor for another factor is different depending on what causal significance the other factors have for the third factor. But causal significance of a factor for the third factor is different, depending on propensities individuals have to locate themselves in subpopulations exemplifying the other factors as types or kinds. The ternary theory too assumes populations where the relations between variables hold. So if we consider a general theory that holds for situation as many as possible, then we certainly need information about causal significance of disjunctive factors for another factor.

My criticism of Hitchcock's arguments for the ternary theory and against the unanimity theory is worth noticing today. Several philosophers of science (Hausman [9], Maslen [15], Northcott [16], Schaffer [18]), which have discussed what should be contrasted as causal relata, argue that the ternary contrast is not only distinct from the binary contrast but also meets the problems with theories of causation allegedly due to the binary contrast. They at least implicitly credit their arguments for the ternary contrast to Hitchcock's argument for the ternary contrast. For example, Hausman's [9] formulation of causal generalizations in some homogeneous background contexts is as follows:

"In population  $P$ ,  $X = x^*$  as compared to  $X = x'$  causes  $Y$ " is true if and only if

- (a) in population  $P$ ,  $Pr(Y | X = x^*) > Pr(Y | X = x')$  and
- (b) the probability difference in (a) is due to the causal inference of  $X = x^*$  as compared to  $X = x'$  in some causally homogeneous circumstance occupied by members of population  $P$ .

Hausman's formulation of causal generalizations is

different from the unanimity theory in two ways. First, a factors  $X$  is replaced by a random variable  $X = x$ . Second, the comparison of a factor  $X$  and  $\sim X$  (i.e., absence of  $X$ ) does not exist. Instead, a value (e.g., smoking a pack of cigarettes) of a random variable  $X$  is compared with the other values (e.g., smoking two packs of cigarettes) of the variable  $X$ . Hitchcock [10] calls this relation of contrast the *ternary* relation. As Hausman [8], [9] mentioned elsewhere, the introduction of random variables and the comparison between values of a variable are indebted to Hitchcock's [10] ternary probabilistic theory of causation. As Hitchcock argues, so does Hausman would claim that the two features of the ternary theory render his formulation of causal generalizations more practical, or practically more competent than causal generalizations founded on the unanimity theory. It is obvious that Hitchcock's ternary theory and Hausman's theory of causal generalizations coincide formally and conceptually. I can also argue that Hausman's theory too is founded on the unanimity theory by showing that the unanimity theory conveys the same information Hitchcock's ternary theory is intended to do.

Some computer scientists, for example, Pearl [17], who have worked much on causality in A.I., do not notice the problem with the ternary contrast Hitchcock employs. My discussions will alert them to watch out toying with the ternary contrast. At least as far as causal relevance is relative to population, the binary contrast and the ternary contrast coincide and moreover, the ternary contrast is reducible to the binary contrast.

#### ACKNOWLEDGMENT

Joonsung Kim thanks three great philosophers of science at the University of Wisconsin-Madison, Elliott Sober, Daniel Hausman and especially Ellery Eells, for their meticulous and helpful comments on the previous drafts of this paper. The main point of this paper should be credited with my mentor, Ellery Eells who untimely passed away, 2006.

#### REFERENCES

- [1] N. Cartwright, "Causal Laws and Effective Strategies," *Noûs* 13, pp. 419-437, 1979. Reprinted in N. Cartwright, *How the Laws of Physics Lie*, Clarendon Press, Oxford, 1983.
- [2] N. Cartwright, *Nature's Capacities and their Measurements*. Clarendon Press, 1989.
- [3] J. Dupré, "Probabilistic Causality Emancipated," In P. A. French, T. E. Uehling, H. K. Wettstein (eds.), *Midwest Studies in Philosophy IX: Causation and Causal Theories*, University of Minnesota Press, Minneapolis, 1984, pp. 169-75.
- [4] E. Eells, "Probabilistic Causal Interaction and Disjunctive Causal Factors," In J. Fetzer (eds.), *Probability and Causality: Essays In Honor of Wesley C. Salmon*. D. Reidel, Dordrecht, 1988, pp. 189-209.
- [5] E. Eells, *Probabilistic Causality*, Cambridge University Press, 1991.
- [6] E. Eells and E. Sober, "Probabilistic Causality and the Question of Transitivity," *Philosophy of Science* 50, pp. 35-57, 1983.
- [7] I. Good, "A Causal Calculus (I)-(II)," *British Journal for the Philosophy of Science* 11, pp. 305-318; 12, pp. 43-51, 1961-1962.
- [8] D. Hausman, *Causal Asymmetries*, Cambridge University Press, 1998.
- [9] D. Hausman, "Probabilistic Causation and Practical Causal Generalizations," In E. Eells and J. Fetzer (eds.), *The Place of Probability in Science*, Springer, 2010, pp. 47-64.
- [10] C. Hitchcock, "A Generalized Probabilistic Theory of Causal Relevance," *Synthese* 97, pp. 335-64, 1993.

- [11] P.W. Holland, "Statistics and Causal Inference (in Theory and Methods)," *Journal of the American Statistical Association*, Vol. 81, pp. 945-960, 1986.
- [12] P. Humphreys, *The Chances of Explanation*, Princeton University Press, 1989.
- [13] J. Kim, "Probabilistic Causality: In Defense of the Unanimity Theory," Ph.D. dissertation, Dept. of Philosophy, University of Wisconsin-Madison, 2002.
- [14] J. Kim, "The Inference of Contrast for Probabilistic Theories of Causality," *International Proceedings of Computer Science and Information Technology: 2012 IACSIT Hong Kong Conferences*, Vol. 25, 2012, pp. 34-38.
- [15] C. Maslen, "Causes, Contrasts and the Nontransitivity of Causation," In J. Collins, N. Hall, & N. Hall, & L.A. Paul (eds.), *Causation and Counterfactuals*, Cambridge: MIT Press, 2004, pp.341-357.
- [16] R. Northcott, "Causation and Contrast Class," *Phil. Stud.* 139, pp. 111-123, 2007.
- [17] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.
- [18] J. Schaffer, "Contrastive Causation," *Philosophical Review* 114, pp. 297-328, 2005.
- [19] B. Skyrms, *Causal Necessity*, Yale University Press, 1980.
- [20] E. Sober, *The Nature of Selection*, MIT Press, 1984.
- [21] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, New York: Springer-Verlag, 1993.
- [22] P. Suppes, *A Probabilistic Theory of Causality*, North-Holland Pub, 1970.