

What Should I Believe About What Would Have Been the Case?

Franz Huber

Department of Philosophy

University of Toronto

franz.huber@utoronto.ca

<http://huber.blogs.chass.utoronto.ca/>

penultimate version: please cite the paper in the

Journal of Philosophical Logic

July 7, 2014

Contents

1	Introduction	3
2	Counterfactuals and Testing	4
3	Probabilistic Metaphysics and Epistemology	6
4	Traditional Metaphysics and Epistemology	14
5	Rank-Theoretic Metaphysics and Epistemology	15
6	The Obvious Observation	19
7	Conditional Independence and Consistency	22
8	What I Should Believe About What Would Have Been the Case	30
9	Comparisons	34

Abstract

The question I am addressing in this paper is the following: how is it possible to empirically test, or confirm, counterfactuals? After motivating this question in section 1, I will look at two approaches to counterfactuals, and at how counterfactuals can be empirically tested, or confirmed, if at all, on these accounts in section 2. I will then digress into the philosophy of probability in section 3. The reason for this digression is that I want to use the way observable absolute and relative frequencies, two empirical notions, are used to empirically test, or confirm, hypotheses about objective chances, a metaphysical notion, as a role-model. Specifically, I want to use this probabilistic account of the testing of chance hypotheses as a role-model for the account of the testing of counterfactuals, another metaphysical notion, that I will present in sections 4 to 8. I will conclude by comparing my proposal to one non-probabilistic and one probabilistic alternative in section 9.

1 Introduction

I claim that the following two subjunctive, or counterfactual, conditionals are true: If I had not had coffee this morning, I would have been tired at noon; and if I had had coffee this morning, I would not have been tired at noon. These conditionals are true despite the fact that I had coffee this morning, and despite the fact that I have not been tired at noon. These facts make the first conditional a genuine contrary-to-fact, or counter-factual, conditional. However, in what follows I will refer to all conditional claims in the subjunctive mood as counterfactuals, even if their antecedent and consequent are true, as is the case for the second conditional.

The question I am addressing in this paper is the following: how is it possible to empirically test, or confirm, such counterfactuals? To me this question is of interest in itself. To fellow philosophers this question might be of interest, because they may sometimes rely on counterfactuals in their philosophical theorizing.

Philosophical problems that have been approached in terms of counterfactuals abound. Counterfactual accounts of causation provide an example from metaphysics (Collins & Hall & Paul 2004). Counterfactual accounts of knowledge provide an example from epistemology (Nozick 1981). Counterfactual accounts of dispositions provide an example from the philosophy of science (Mumford 1998). Even if one holds that all these counterfactual approaches are misguided, one might still be interested in finding out under what conditions one ought to believe a given counterfactual to be true, or to be false. Indeed, unless one is an expressivist about counterfactuals and holds that counterfactuals do not express propositions, and do not have truth conditions (Edgington 2008, Spohn 2013), an answer to this question is a precondition for the possibility of criticizing such counterfactual accounts.

The plan for this paper is as follows. I will begin by looking at two approaches to counterfactuals, and by considering how counterfactuals can be empirically tested, or confirmed, if at all, on these two accounts in section 2. In section 3 (and section 4) I will then digress into the philosophy of probability. The reason for this digression is that I want to use the way observable absolute frequencies and observable relative frequencies, two empirical notions, are used to empirically test, or confirm, hypotheses about objective chances, a metaphysical notion, as a role-model. More specifically, I want to use this probabilistic account of the testing of chance hypotheses as a role-model for the account of the testing of counterfactuals, another metaphysical notion, that I will present in sections 4 to 8. I will conclude by comparing my proposal to one non-probabilistic and one probabilistic alternative in section 9.

2 Counterfactuals and Testing

According to one version of the similarity approach to counterfactuals (Stalnaker 1968, Lewis 1973), a counterfactual conditional $\alpha \square\rightarrow \gamma$ is true at a possible world $w \in W$ in a model $\mathcal{M} = (W, (\$w)_{w \in W}, \llbracket \cdot \rrbracket)$ if and only if all of the most $\$w$ -similar $\llbracket \alpha \rrbracket$ -worlds are $\llbracket \gamma \rrbracket$ -worlds. Besides the system of spheres $(\$w)_{w \in W}$ and the evaluation function $\llbracket \cdot \rrbracket$ these models contain a set of possible worlds W . I do not know what possible worlds are. However, whatever possible worlds are, I know that they are *exclusive* in the sense that *at most one* element of W is actual or designated. For present purposes this is all we need to know about them.

One way to think of, or construct, the set of possible worlds W is as the Cartesian product $\times_{X \in \mathcal{X}} R(X)$ of the ranges, or sets of possible values, $R(X)$ of a set \mathcal{X} of *singular* variables X . This way of thinking of, or constructing, possible worlds is popular in the literature on causal modeling. In the philosophical literature the set of possible worlds is often constructed from the underlying formal language. This is effectively the same thing, except that the singular variables are binary propositional variables. In some treatments of modal logic the set of possible worlds is assumed to be given as a primitive.

The similarity approach to counterfactuals gives us a precise semantics for counterfactuals. However, at least on the versions of Stalnaker (1968) and Lewis (1973; 1979), the similarity approach does not tell us how to empirically test, or confirm, counterfactuals. According to Stalnaker (1968: 177ff) “[i]t is because counterfactuals are generally about possible worlds which are very much like the actual one, and defined in terms of it, that evidence is so often relevant to their truth”. This is simply too thin an account of the testing of counterfactuals to be evaluated. Lewis (1979: 465) admits that (something like) the above definition “does little to predict the truth values of particular counterfactuals in particular contexts” and adds his “system of weights or priorities” (Lewis 1979: 472). However, he then reverses the order of analysis: “we must use what we know about counterfactuals to find out about the appropriate similarity relation—not the other way around” (Lewis 1979: 467). This reverse engineering does little to make counterfactuals empirically testable. Indeed, if it were part of the analysis itself, rather than descriptive of the methodology of analysis on the meta-level, it would render counterfactuals untestable and, in this sense, without empirical content.

A more recent approach to counterfactuals explicitly rejects their untestability, ties them to interventions, and focuses on counterfactuals of the form: $F = f \square\rightarrow G = g$, where F and G are *generic* variables. This is the interventionist approach to causal counterfactuals (Woodward 2003).

According to Woodward (2003: 72-73) “one way ... of testing ... counterfactual claims ... will be to carry out the interventions described in their antecedents and then check to see whether certain correlations hold.” As Woodward explains, he thinks a sample provides confirmation for a counterfactual of the form ‘if the value of F were set to f , then the value of G would change to g ’ if there is perfect correlation in the sense that all individuals in the sample, all individuals whose value for F has been set to f , take on the value g for G . If only some, but not all individuals in the sample take on the value g for G , and their relative frequency is different from the unconditional probability, or the unconditional chance, that an arbitrary or “randomly chosen” individual has the value g for G , then the sample provides confirmation for a counterfactual of the form ‘if the value of F were set to f , then the probability or chance that the value of G is g would change to p ’, where p is close to the observed relative frequency in the sample.

The interventionist approach to causal counterfactuals tells us, more or less¹, how to empirically test, or confirm, counterfactuals. However, it does not give us a precise semantics for counterfactuals.²

A problem arises. It is unclear how, if not impossible, to combine the nice feature of the similarity approach, viz. its precise semantics, with the nice feature of the interventionist approach, viz. its account of testing. The reason is that a precise semantics is possible on the similarity approach, only because the latter relies on singular variables. The particular account of testing sketched above is possible on the interventionist approach, only because the latter relies on generic variables. And singular variables and generic variables, while mathematically indistinguishable – both are measurable functions – are only superficially similar entities.

Singular variables are functions from a set of possible worlds W into a range R . They generalize the more familiar notion of a binary propositional variable. Singular variables are mainly used in philosophy, especially if the concern is the analysis or explication of token or actual causation between events (Baumgartner & Glynn 2013) – or, more cautiously, the relation of actual causation (Paul 2000).

¹On the account of testing envisaged by Woodward (2003) and sketched above one cannot distinguish between confirmation for a counterfactual of the form ‘if the value of F were set to f , then the value of G would change to g ’ and confirmation for a counterfactual of the form ‘if the value of F were set to f , then the probability or chance that the value of G is g would change to p ’.

²Briggs (2012), Galles & Pearl (1998), Halpern (2013), Zhang (2013), and Zhang & Lam & De Clercq (2013) give us a precise semantics for counterfactuals. However, they all seem to understand the variables in a singular sense. Therefore Woodward’s (2003) account of testing is not available to them.

For instance, the singular variable $X : W \rightarrow \mathbb{N}$ may assign the value m to a possible world, if m mg of the pain killer ibuprofen are administered to Simone de Beauvoir in the morning of January 1, 1950, in this possible world. It is because X is defined on an *exclusive* set of possible worlds W that a precise semantics is possible.

In contrast to this, generic variables are functions from a population or domain of individuals from which one can draw samples, D , into a range R . They are mainly used in the natural and social sciences and other areas where statistics is applied, especially if the concern is the investigation of causal relevance relations between properties.

For instance, the population may be the set of people of a certain age and in a certain geographical region. The generic variable $F : D \rightarrow \mathbb{N}$ may assign the value m to an individual from this population, if m mg of the pain killer ibuprofen are administered to this individual. It is because F is defined on a *non-exclusive* domain D from which one can draw samples that the account of testing sketched above is possible.

The task that arises from this problem is two-fold. First, we have to come up with an account that gives us a precise semantics for counterfactuals, as the similarity approach does. Second, we have to state conditions under which these precisely defined counterfactuals, and not some generic type-level substitute, can be empirically tested, or confirmed. In order to cope with this task let us first digress into the philosophy of probability and look for a role-model.

3 Probabilistic Metaphysics and Epistemology

In probabilistic epistemology there is the notion of a subject S 's credence in a hypothesis H given various data E , $Cr_S(H | E)$. In probabilistic metaphysics there is the notion of the objective chance at a certain world w and at a certain time t that a proposition A is true, or that an event a takes place, $ch_{w,t}(A)$. In addition there is the empirical notion of the observable absolute frequency, and the empirical notion of the observable relative frequency, of a certain outcome x , or an event $X = x$, in the first n repetitions of some experiment, or type of event, X in the actual world w , $relfr(x, X^n(w))$. Here $X = x$ is the proposition $\{w : X(w) = x\}$.

These three probabilistic notions – subjective credences, objective chances, and observable relative frequencies – are linked by three principles (Spohn 2010). The **Principal Principle** (Lewis 1980) relates subjective credences and objective chances. The **Strong Law of Large Numbers** (Earman 1992 and many others)

relates objective chances and limiting relative frequencies (for independent and identically distributed random variables). The Straight Rule (Reichenbach 1939) relates relative frequencies and subjective credences. A limiting version of the Straight Rule, the Axiom of Convergence or Reichenbach Axiom (Carnap 1980: §20), relates limiting relative frequencies and limiting subjective credences. The latter principles will play no role in this paper. However, the Principal Principle and the Strong Law of Large Numbers will play a role. Together they imply that observable relative frequencies raise and lower subjective credences in hypotheses about objective chances. This will be the probabilistic role-model for the account of the testing of counterfactuals that I will present in sections 4 to 8.

The **Principal Principle** says that a subject S 's a priori credence in some proposition A given that the objective chance of A at time t equals x – as well as, perhaps, information E_t that is admissible at t , but no further information that is inadmissible at t – should equal x :

$$Cr_S(A \mid ch_t(A) = x \cap E_t) = x.$$

Lewis (1980) argues that the complete history of world w up to time t , $H_{w,t}$, is admissible at t . He argues further that the world w 's theory of indeterministic alethic modality, or chance, T_w is admissible at any time. This allows him to derive the following reformulation of the Principal Principle:

$$ch_{w,t}(A) = Cr_S(A \mid H_{w,t} \cap T_w).$$

The reformulation says that the chance distribution of a world w at a time t comes from a subject S 's a priori credence function by conditionalizing on the complete history of world w up to time t and the world w 's theory of indeterministic alethic modality, or chance, T_w . For an excellent discussion of the Principal Principle see Briggs (2009).

The statement of the Strong Law of Large Numbers requires a little bit of terminology. We start with a non-empty set of outcomes or *possibilities* Ω and a field or an algebra $\mathcal{A} \subseteq \wp(\Omega)$ of events or *propositions* $A \subseteq \Omega$ over Ω . Here \mathcal{A} is a finitary/ σ -/ γ - algebra if and only if (1) $\emptyset \in \mathcal{A}$, (2) $\overline{A} \in \mathcal{A}$ if $A \in \mathcal{A}$, and (3) $\bigcup \mathcal{B} \in \mathcal{A}$ if each $B \in \mathcal{A}$ for each $B \in \mathcal{B}$ and all finite/countable/arbitrary $\mathcal{B} \subseteq \wp(\Omega)$. (We will use γ -algebras in section 5.) (Ω, \mathcal{A}) is a *measurable space*. We define a probability measure \Pr on \mathcal{A} to obtain a *probability space* $(\Omega, \mathcal{A}, \Pr)$.

A function $\Pr : \mathcal{A} \rightarrow [0, 1]$ is a finitely additive/countably (or σ -) additive *probability measure* on (Ω, \mathcal{A}) just in case (P1) $\Pr(\Omega) = 1$ and (P2) $\Pr(\bigcup_{B \in \mathcal{B}} B) = \sum_{B \in \mathcal{B}} \Pr(B)$ for all finite/countable collections \mathcal{B} of mutually exclusive $B \in \mathcal{A}$.

Let (Υ, \mathcal{V}) be another measurable space. A function $X : \Omega \rightarrow \Upsilon$ is \mathcal{A} - \mathcal{V} -measurable just in case for all $V \in \mathcal{V}$: $X^{-1}(V) \in \mathcal{A}$. An \mathcal{A} - \mathcal{V} -measurable function X is sometimes referred to as a *random variable*.

The algebra $\sigma(X) \subseteq \mathcal{A}$ generated by the \mathcal{A} - \mathcal{V} -measurable function $X : \Omega \rightarrow \Upsilon$ is the smallest σ -algebra containing $X^{-1}(V)$ for all $V \in \mathcal{V}$. The algebra $\gamma(X) \subseteq \mathcal{A}$ generated by the \mathcal{A} - \mathcal{V} -measurable function $X : \Omega \rightarrow \Upsilon$ is the smallest γ -algebra containing $X^{-1}(V)$ for all $V \in \mathcal{V}$. (We will use $\gamma(X)$ in section 5.)

A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *independent* in the sense of \Pr just in case for any finite subfamily X_{i_1}, \dots, X_{i_n} and all $A_{i_1} \in \sigma(X_{i_1}), \dots, A_{i_n} \in \sigma(X_{i_n})$:

$$\Pr(A_{i_1} \cap \dots \cap A_{i_n}) = \prod_{j=1, \dots, n} \Pr(A_{i_j}).$$

A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *identically distributed* in the sense of \Pr just in case for all $i, j \in I$ and all $V \in \mathcal{V}$:

$$\Pr(X_i^{-1}(V)) = \Pr(X_j^{-1}(V)).$$

Consider a sequence of random variables X_1, \dots, X_n, \dots . Think of the X_i as repetitions of an experiment type, or as the single infinite experiment $X^\infty : \Omega \rightarrow \Upsilon^\infty$ from the set of outcomes or possibilities Ω into the infinite repetition Υ^∞ of the set of values Υ . One may think of Ω as the set of possible values of this infinite experiment X^∞ , i.e. $\Omega = \Upsilon^\infty$. The *relative frequency* of value $v \in \Upsilon$ in possibility $\omega \in \Omega$ in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$ compares v to *all its competitors* $v' \in \Upsilon$. It is the proportion of the number of v -outcomes among the sum of the number of all v' -outcomes:

$$|\{i : X_i(\omega) = v, 1 \leq i \leq n\}| / \sum_{v' \in \Upsilon} |\{i : X_i(\omega) = v', 1 \leq i \leq n\}|.$$

The complicated sum in the denominator is always equal to n . The reason for stating the definition of relative frequencies in this complicated way is that this makes clear that, when forming the relative frequency of an outcome, we compare the number of times the outcome has occurred to the sum of the number of times some outcome or other has occurred. That is, when we evaluate the performance of the outcome of interest, we compare the performance of this outcome to the performance of *all* its competitors. This is one way of evaluating the performance of the outcome of interest. It is not the only way, though. Later on, when it comes to testing counterfactuals rather than chance hypotheses, I will propose a slightly different way of evaluating the performance of the outcome of interest.

Let $afr(v, X^n(\omega)) = |\{i : X_i(\omega) = v, 1 \leq i \leq n\}|$ be the *absolute frequency* of value v in possibility ω in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$. Then the relative frequency can be written as:

$$relfr(v, X^n(\omega)) = afr(v, X^n(\omega)) / \sum_{v' \in \Upsilon} afr(v', X^n(\omega))$$

The relative frequency of the set of values $V \in \mathcal{V}$ in possibility ω in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$ is defined as follows:

$$relfr(V, X^n(\omega)) = \sum_{v \in V} relfr(v, X^n(\omega)).$$

Now we are in a position to state the **Strong Law of Large Numbers**. For each countably additive probability measure \Pr there is an $A \in \mathcal{A}$ with $\Pr(A) = 1$ such that for all $\omega \in A$: the relative frequency of v in ω in $X_1(\omega), \dots, X_n(\omega)$ converges to the probability of v in ω in the sense that

$$\lim_{n \rightarrow \infty} |relfr(v, X^n(\omega)) - \Pr(\{\omega' \in \Omega : X(\omega') = v\})| = 0,$$

where X_1, \dots, X_n, \dots is a sequence of independent and identically distributed \mathcal{A} - \mathcal{V} -measurable functions and X is an arbitrary one of them.

If, and only if, Υ is a set of numbers so that we can form the average or *mean* of X in $X_1(\omega), \dots, X_n(\omega)$, $\sum_{i=1, \dots, n} X_i(\omega) / n$, and the *Pr-expected value* of X , $\sum_{v \in \Upsilon} v \cdot \Pr(\{\omega' \in \Omega : X(\omega') = v\})$, where, for the sake of simplicity, Υ is assumed to be countable, the Strong Law of Large Numbers says that the mean converges to the expected value in all worlds except, perhaps, those in a set of measure zero.

It is important to note that the \mathcal{A} - \mathcal{V} -measurable functions $X_i : \Omega \rightarrow \Upsilon$ are what we have called *singular variables* above. They are defined on an exclusive set of possibilities Ω . X_1 may assign value 1 to a possible world if Simone de Beauvoir feels relief of pain in the late morning of January 1, 1950, in this possible world, and 0 otherwise. X_2 may assign value 1 to a possible world if Jean-Paul Sartre feels relief of pain in the late morning of January 1, 1950, in this possible world, and 0 otherwise. X_3 may assign value 1 to a possible world if Albert Camus feels relief of pain in the late morning of January 1, 1950, in this possible world, and 0 otherwise. And so on. If we have a sequence of such singular variables X_1, \dots, X_n, \dots , we can use this sequence to construct a population $D = \{X_i : i \in \mathbb{N}\}$. We may think of the function $G : D \rightarrow \Upsilon$ as the generic variable

that, roughly, assigns value 1 to an individual, Simone de Beauvoir or Jean-Paul Sartre or Albert Camus or Gabriel Marcel, if this individual feels relief of pain in the late morning of January 1, 1950, and value 0 otherwise.

That the sequence of singular variables X_1, \dots, X_n, \dots is independent means that whether or not X_1 takes on some value x_1 does not affect the probability that X_2 takes on some value x_2 . Whether or not Simone de Beauvoir feels relief of pain in the late morning of on January 1, 1950, does not affect the probability that Jean-Paul Sartre feels relief of pain in the late morning of January 1, 1950. That the sequence of singular variables X_1, \dots, X_n, \dots is identically distributed means that the probability that X_1 takes on the value x is identical to the probability that X_3 takes on this value x . The probability that Simone de Beauvoir feels relief of pain in the late morning of January 1, 1950, is identical to the probability that Albert Camus feels relief of pain in the late morning of January 1, 1950. These are strong assumptions.

In contrast to the Principal Principle, which is a normative constraint on a priori credence functions, the Strong Law of Large Numbers is a theorem of the probability calculus. As a theorem it holds independently of how we interpret the countably additive probability measure Pr . Therefore we can simply choose to read it as relating (limiting) observable relative frequencies and objective chances.

For each objective chance space $\mathbb{C} = (\Omega, \mathcal{A}, ch)$ with the objective chance measure ch being countably additive, each measurable space $(\mathcal{Y}, \mathcal{V})$, and each sequence of \mathcal{A} - \mathcal{V} -measurable functions X_1, \dots, X_n, \dots which is independent and identically distributed in the sense of ch there is an $A \in \mathcal{A}$ with $ch(A) = 1$ such that for all $\omega \in A$: the observable relative frequency of v in ω in $X_1(\omega), \dots, X_n(\omega)$ converges to the objective chance of v in ω in the sense that

$$\lim_{n \rightarrow \infty} \left| \text{relfr}(v, X^n(\omega)) - ch(\{\omega' \in \Omega : X(\omega') = v\}) \right| = 0.$$

Note that the sequence of random variables X_1, \dots, X_n, \dots is assumed to be independent and identically distributed in the sense of the objective chance measure ch . Presumably it is more difficult to establish that a sequence of experiments, or a sequence of repetitions of an experiment type, is independent and identically distributed in the sense of the objective chance measure of the actual world than to establish that it is so in the sense of a subject S 's credence function. Conversely, once these two assumptions are established for the objective chance measure of the actual world, the Strong Law of Large Numbers says something about the objective chances of the actual world rather than some subject S 's credences.

The point of importance for the present task is that the Principal Principle implies that observable relative frequencies raise and lower subjective credences

in hypotheses about objective chances. More specifically, suppose a subject S 's a priori credence function Cr_S satisfies the Principal Principle. Then learning, or conditionalizing on, information about the observable relative frequencies of sequences of random variables X_1, \dots, X_n, \dots that are independent and identically distributed in the sense of the objective chance measure ch affects the subject S 's credences in hypotheses about the objective chance measure ch :

$$\begin{aligned}
Cr_S(ch(X = x) = p \mid relfr(x, X^n) = q) &= Cr_S(relfr(x, X^n) = q \mid ch(X = x) = p) \cdot \\
&\quad \frac{Cr_S(ch(X = x) = p)}{Cr_S(relfr(x, X^n) = q)} \\
&= \frac{n! \cdot p^{n \cdot q} \cdot (1 - p)^{n \cdot (1 - q)}}{(n \cdot q)! \cdot (n \cdot (1 - q))!} \cdot \\
&\quad \frac{Cr_S(ch(X = x) = p)}{Cr_S(relfr(x, X^n) = q)} \\
&\propto f(p, q, n) \cdot Cr_S(ch(X = x) = p)
\end{aligned}$$

The first equation results from an application of Bayes' Theorem. The second equation results from an application of the Principal Principle. All this, and more, is explained much better in Lewis (1980: 282-287).

This means the following. The subject S 's posterior credence in the hypothesis that the objective chance of outcome x in experiment X equals p conditional on the information that the observable relative frequency of outcome x in the first n repetitions of the experiment X equals q , $Cr_S(ch(X = x) = p \mid relfr(x, X^n) = q)$, is proportional to (\propto) the product of the subject S 's prior credence in this chance hypothesis, $Cr_S(ch(X = x) = p)$, and a parameter $f(p, q, n)$ that can be computed from the hypothesized objective chance p of the outcome x , the observable relative frequency q , and the number n of repetitions of the experiment X .

For any n , the empirically computable parameter $f(p, q, n)$ is the higher, the closer the observable relative frequency q is to the hypothesized objective chance p . The curve representing this relationship is the steeper, the greater n . The Strong Law of Large Numbers then implies that the objective chance is maximal ($= 1$) that the subject S 's posterior, or conditional, credence

$$Cr_S(ch(X = x) = p \mid relfr(x, X^n) = q)$$

eventually is the highest for this chance hypothesis $ch(X = x) = p$ that is true in the actual world. At least, this is so provided that $Cr_S(ch(X = x) = p)$ and $Cr_S(relfr(x, X^n) = q)$, for any n , are not minimal ($= 0$) to begin with. Again, all this, and more, is explained much better in Lewis (1980: 282-287).³

Ignoring complications that will be spelled out in detail later on when we turn to the testing of counterfactuals, this equation allows us to learn about the objective chance that a coin lands heads on a particular toss, and whether the administration of 300 mg of ibuprofen increases the chance that pain is relieved shortly afterwards.

More specifically, we can use the observable relative frequency of heads in a sequence of independent and identically distributed tosses of a coin to learn about the objective chance that the coin will land heads on the next toss. We make some assumptions about the objective chances, viz. that the coin's landing heads on the first toss does not affect the chance that it lands heads on the second toss, as well as that the chance that the coin lands heads on the first toss is identical to the chance that it lands heads on the second toss. Or rather, we design our experiments in such a way that these two assumptions are met. Given these two assumptions we can then use the empirically accessible information about the observable relative frequency of heads to infer, with objective chance one and in the limit, what the objective chance is that the coin lands heads on the next toss.

In a similar way we can use the observable relative frequency of relief of pain in a sequence of independent and identically distributed experiments where subjects are administered 300 mg of ibuprofen to learn whether the administration of 300 mg of ibuprofen increases the chance that pain is relieved shortly afterwards. Again, we make some assumptions about the objective chances. First, we assume that Simone de Beauvoir's feeling relief of pain does not affect the chance that Jean-Paul Sartre feels relief of pain (given that both are administered 300 mg of ibuprofen slightly earlier). Second, we assume that the chance that Simone de Beauvoir feels relief of pain (given that she is administered 300 mg of ibuprofen slightly earlier) is identical to the chance that Jean-Paul Sartre feels relief of pain (given that he is administered 300 mg of ibuprofen slightly earlier). Or rather, we design our experiments in such a way that these assumptions are met. Given these assumptions we can then use the empirically accessible information about the observable relative frequency of relief of pain to infer, with objective chance one (conditional on the subjects being administered 300 mg of ibuprofen) and in the limit, whether the administration of 300 mg of ibuprofen to Albert Camus increases the chance that he feels relief of pain shortly afterwards.

³The Straight Rule, and hence the Axiom of Converge or Reichenbach Axiom, does not enable us to learn hypotheses about objective chances. As is well known, though, it can be reformulated as a rule for inferring limiting relative frequencies. Then one can show that it eventually conjectures values that are arbitrarily close to the limiting relative frequencies in the actual world, provided the latter exist.

This probabilistic model of the testing of chance hypotheses is the role-model for the account of the testing of counterfactuals that I will present in sections 4 to 8. It allows us to characterize our task more precisely.

First we have to find (i) appropriate doxastic states $Cr_S^*(H | E)$ of conditional form – otherwise, without conditional form, we cannot learn, or conditionalize. Next we have to find (ii) appropriate conditions of counterfactual independence and counterfactual identical distribution (rather than independence and identical distribution in the sense of the objective chance measure ch) for sequences of experiments X_1, \dots, X_n, \dots . And then we have to find (iii) appropriate observable properties $relfr^*$ of these counterfactually independent and identically distributed sequences of experiments X_1, \dots, X_n, \dots .

Second we have to find a principle analogous to the Principal Principle that entails that, for some appropriate functions f^* and g :

$$Cr_S^*(X = x \square \rightarrow Y = y | relfr^* = q^*) \propto^* g(f^*(p^*, q^*, n), Cr_S^*(X = x \square \rightarrow Y = y))$$

Here p^* is the hypothesized counterfactual distance to the closest $X = x$ -worlds. The latter is the counterfactual analogon of the hypothesized objective chance p of the outcome x . It is implicit in the counterfactual $X = x \square \rightarrow Y = y$. The idea is that, roughly, the subject S 's posterior doxastic attitude towards a counterfactual should come from, or be appropriately related to (\propto^*), her prior doxastic attitude towards this counterfactual as well as empirically accessible information.

In a final step we have to prove a theorem analogous to the Strong Law of Large Numbers that entails that learning, or conditionalizing on, information about the appropriate observable properties $relfr^*$ of counterfactually independent and identically distributed sequences of experiments X_1, \dots, X_n, \dots successfully affects the subject S 's doxastic attitude towards counterfactuals $X = y \square \rightarrow Y = y$. Here successfully affecting the subject S 's attitude towards counterfactuals means that the subject eventually comes to believe every true counterfactual, and that she eventually becomes to disbelieve every false counterfactual, to the extent that this is possible (we will see that, without further assumptions, the ranking calculus and the Royal Rule merely imply that eventually all disbelieved counterfactuals are false).

This ends our digression into the philosophy of probability. Let us return to more familiar territory.

4 Traditional Metaphysics and Epistemology

In traditional epistemology there is the notion of a subject S 's conditional belief $Bel_S(H | E)$ in hypothesis H given various data E . In traditional metaphysics there is the notion of the counterfactual dependence of certain propositions C , or events c , on certain propositions A , or events a , $\pm A \Box \rightarrow \pm C$. In addition there is the empirical notion of what one observes to happen, and what one observes to fail to happen.

One option for coping with our task, parsimonious both epistemologically and ontologically, is to *probabilify* these three traditional notions of conditional belief, counterfactual dependence, and observables. One way of probabilifying works by *replacement*: replace a subject S 's conditional belief $Bel_S(H | E)$ in hypothesis H given various data E by the subject S 's credence $Cr_S(H | E)$ in this hypothesis H given those data E along the lines of Jeffrey (1970). Replace counterfactual dependence of certain propositions C on certain propositions A , $\pm A \Box \rightarrow \pm C$, by the objective chances of these propositions $\pm C$ given those propositions $\pm A$ at the actual world at a certain past time t , $ch_{w,t}(\pm C | \pm A)$, along the lines of Edgington (2008). And replace what one observes to happen, and to fail to happen, by a report of these observables in the form of relative frequencies.

A different way of probabilifying works by *reduction*: reduce a subject S 's conditional belief $Bel_S(H | E)$ in hypothesis H given various data E to the subject S 's credence $Cr_S(H | E)$ in this hypothesis H given those data E along the lines of Leitgeb (2013). Reduce counterfactual dependence of certain propositions C on certain propositions A , $\pm A \Box \rightarrow \pm C$, to the sets of possible worlds w where the objective chances of these propositions $\pm C$ given those propositions $\pm A$ are 1, $ch_w(\pm C | \pm A) = 1$, along the lines of Leitgeb (2012a; b) – or else, do something analogous along the multidimensional lines of Bradley (2012). And reduce what one observes to happen, and to fail to happen, to a report of these observables in the form of relative frequencies.

I do not believe in replacement, and I am skeptical of reduction because of the negative results of Lin & Kelly (2012). More specifically, Leitgeb (2013) proposes that an agent believes a proposition B if and only if there is a proposition C implying B such that the agent's subjective probabilities for C conditional on any A consistent with C are above a certain threshold that is not smaller than $1/2$. This notion of belief satisfies the AGM axioms of belief revision (Alchourrón & Gärdenfors & Makinson (1985). However, as Lin & Kelly (2012) show, there is no "sensible" belief revision method that tracks conditionalization and satisfies these AGM axioms. This means that what an agent ends up believing according

to Leitgeb (2013), and any other sensible proposals satisfying the AGM axioms, if she first conditionalizes her subjective probabilities on evidence E and then extracts her beliefs will, in general, not coincide with what she ends up believing if she first extracts her beliefs from her subjective probabilities and then revises those beliefs by evidence E .

Perhaps my skepticism is misguided. Even if this were the case, the result of reducing conditional belief and counterfactual dependence to probabilistic notions would still be a different one than the result of taking the notions of conditional belief and counterfactual dependence at face-value and formalizing them in terms of ranking functions, as I will do in the next section. I will turn to a comparison of the two results in section 9, once we have established the second one in sections 5 to 8.

5 Rank-Theoretic Metaphysics and Epistemology

A function $R : \mathcal{A} \rightarrow \mathbb{N} \cup \{\infty\}$ from an algebra of propositions $\mathcal{A} \subseteq \wp(W)$ over a set of possible worlds W into the natural numbers \mathbb{N} extended by ∞ is a *ranking function* just in case the tautological proposition W is assigned rank 0, $R(W) = 0$, the contradictory proposition \emptyset is assigned rank ∞ , $R(\emptyset) = \infty$, and the rank of a disjunction $A \cup B$ is the minimum of the ranks of the disjuncts A, B , $R(A \cup B) = \min\{R(A), R(B)\}$. Conditional ranks are defined as differences of unconditional ranks: $R(A | B) = R(A \cap B) - R(B)$, provided $R(B) < \infty$.

Spohn (1988) introduces ranking functions to represent conditional belief. He interprets the numbers or ranks doxastically as grades of disbelief. A proposition A is disbelieved just in case its rank is positive, $R(A) > 0$. A proposition A is believed just in case its negation \bar{A} is disbelieved, $R(\bar{A}) > 0$. The first constraint says one should not disbelieve the tautological proposition. The second constraint says that one should disbelieve the contradictory proposition to the highest degree possible. Given the definition of conditional ranks the second constraint says one should disbelieve the contradictory proposition conditional on every proposition with a finite rank. Part of what the third constraint says is that one should disbelieve a disjunction just in case one disbelieves all its disjuncts. This requirement is then extended to conditional beliefs. Given the definition of conditional ranks the third constraint says that one should conditionally disbelieve a disjunction just in case one conditionally disbelieves all its disjuncts. The surprising thing is that we need numbers, ranks, to deal with the qualitative notion of conditional belief. The thing to get used to is that high numbers represent *disbeliefs*.

Perhaps it is helpful to compare probability theory and ranking theory with different sports games. In some sports games such as ice hockey the total score of a team is the *sum* of the individual scores: sum over the goals scored in the first and second and third period. In other sports games such as (high) jumping the total score of an athlete is the *best* individual score: the (high) jump that counts is the furthest (or highest) jump. Evaluating sports games in the first way is akin to probabilistic reasoning: the probability of a proposition is the sum of the probabilities of the possible worlds that comprise the proposition. Evaluating sports games according to the second way is akin to rank-theoretic reasoning: the rank of a proposition is the minimum of the ranks – the lowest or best rank – of the possible worlds that comprise the proposition. Just as both ice hockey and (high) jumping are olympic disciplines, both probability theory and ranking theory are formal tools that can be fruitfully applied to tackle philosophical problems.

Let us do just that. In rank-theoretic epistemology there is the notion of a subject S 's grade of disbelief $R_S(H | E)$ in hypothesis H given various data E . In rank-theoretic metaphysics there is the notion of the counterfactual distance $r_{w,c}(A)$ of a proposition A , or an event a , from certain worlds w in certain contexts c .⁴ In addition there is the empirical notion of the observable absolute failure, $af(x, X^n(w))$, and there is the empirical notion of the observable relative failure, $rf(x, X^n(w))$, of certain outcomes x , or events $X = x$, in the first n repetitions of some experiment, or type of event, X in the actual world w .

I propose to complement the probabilistic metaphysics and epistemology from section 3 with rank-theoretic accounts of the traditional and familiar notions of conditional belief, counterfactual dependence, and observables. To this end I will first briefly explain why rank-theoretic grades of disbelief $R_S(H | E)$ provide an adequate account of conditional belief. Then I will sketch why rank-theoretic counterfactual distances $r_{w,c}(A)$ provide an adequate account of counterfactual dependence. In a third step I will introduce the new rank-theoretic notions of observable absolute failure and the observable relative failure that will allow us to empirically test, or confirm, counterfactuals.

The first argument for the thesis that ranking functions provide an adequate account of conditional belief is due to Hild & Spohn (2008). It is deontological in spirit. Ranking functions induce belief sets that are consistent and deductively closed as already required by Hintikka (1961). In addition ranking functions guide the revision of the induced belief sets in accordance with the AGM constraints on

⁴I will mostly ignore contexts, as the context sensitivity of counterfactuals does not play a role for the purposes of this paper.

belief revision (Alchourrón & Gärdenfors & Makinson 1985). However, these AGM constraints do not say anything about *iterated* belief revisions. Therefore several authors – Boutilier (1996), Darwiche & Pearl (1997), and many others (see Huber 2013a; b) – have proposed various constraints in addition to the AGM constraints to guide iterated belief revisions. These additional constraints are also satisfied by ranking functions (Spohn 2012: §5.6). However, the latter go beyond these constraints. Hild & Spohn (2008) argue that one has to go all the way to ranking functions in order to adequately deal with iterated belief revisions by proving a representation theorem to the effect that, roughly, all and only ranking functions represent iterated belief revisions of a “reasonable” sort. Very roughly, all and only ranking functions obey the duties of iterated belief revision.

The second argument (discussed in Brössel & Eder & Huber 2013) provides a means-ends justification in the instrumental spirit of epistemic consequentialism (Percival 2002, Stalnaker 2002). Besides the above mentioned synchronic rules for organizing one’s rank-theoretic grades of disbelief at a given moment in time ranking theory also includes diachronic rules for updating one’s rank-theoretic grades of disbelief across time if new information of various formats is received. One rule is defined for the case where the new information comes in form of a proposition the subject becomes certain of between the earlier and the later time. It requires the subject’s new ranking function to equal her old ranking function conditional on the information received. Another rule is defined for the case where the new information comes in form of a partition of the space of possibilities plus numbers characterizing the new grades of disbelief in the elements of the partition (for this case one uses Jeffrey conditionalization in probability theory). Yet another rule is defined for the case where the new information comes in form of the degrees by which the grades of disbelief in the elements of such a partition have changed between the earlier and the later time (for this case one uses Field conditionalization in probability theory).

Elsewhere (Huber 2007) I show that a subject’s belief set is, and will always be, consistent and deductively closed, possibly conditional on some element of the given partition, just in case the subject satisfies the synchronic and diachronic rules of ranking theory. That is, obeying the normative constraints of the ranking calculus is a provably necessary and sufficient means to attaining the end of being “diachronically consistent and closed.” The latter is a provably necessary, but insufficient, means to attaining the end of always having only true beliefs, and as many as possible thereof. Therefore, to the extent that one aims at having only true beliefs, and as many as possible thereof, one should obey the synchronic and diachronic rules of the ranking calculus.

Rank-theoretic counterfactual distances are introduced in Huber (2014) as new foundations for counterfactuals. One way to think of this approach is as a variant of the Stalnaker-Lewis approach in terms of similarity. The variation concerns the talk of similarity to a possible world w . The latter is represented by the values of a selection function for w (Stalnaker 1968) or by a sphere of similarity around w (Lewis 1973). I replace this talk by the talk of the counterfactual distance from w , which is represented by the objective ranking function of w .

Formally, I substitute a family of objective ranking functions, one for each possible world, for a selection function or for a system of spheres of similarity. Materially, I treat counterfactual distance as an undefined primitive. However, as I will explained shortly, under certain assumptions hypotheses about counterfactual distances can be empirically tested, just as hypotheses about objective chances can be empirically tested under certain assumptions.

These objective ranking functions which represent counterfactual distances provide an adequate account of counterfactual dependence. They do so in at least the following minimal sense of adequacy. Define $\alpha \square \rightarrow \gamma$ to be true at a world $w \in W$ in a model $\mathcal{M} = (W, (r_w)_{w \in W}, \llbracket \cdot \rrbracket)$ just in case all of the r_w -closest $\llbracket \alpha \rrbracket$ -worlds are $\llbracket \gamma \rrbracket$ -worlds. Then the basic conditional logic \mathbf{V} is sound and complete, or adequate, with respect to this semantics.

We need a little bit more terminology before we turn to the new notions of observable absolute and relative failures. A function $R : \mathcal{A} \rightarrow \mathbb{N} \cup \{\infty\}$ is a finitely/countably/completely minimitive *ranking function* on a measurable space (Ω, \mathcal{A}) just in case (R1) $R(\emptyset) = \infty$ and $R(\Omega) = 0$ and (R2) $R(\bigcup_{B \in \mathcal{B}} B) = \min \{R(B) : B \in \mathcal{B}\}$ for all finite/countable/arbitrary collections \mathcal{B} of $B \in \mathcal{A}$. We define a ranking function R on \mathcal{A} to obtain a *ranking space* (Ω, \mathcal{A}, R) .

Let (Υ, \mathcal{V}) be another measurable space and recall that a function $X : \Omega \rightarrow \Upsilon$ is \mathcal{A} - \mathcal{V} -measurable just in case for all $V \in \mathcal{V}$: $X^{-1}(V) \in \mathcal{A}$. A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *independent* in the sense of R just in case for any finite subfamily X_{i_1}, \dots, X_{i_n} and all $A_{i_1} \in \mathcal{V}(X_{i_1}), \dots, A_{i_n} \in \mathcal{V}(X_{i_n})$:

$$R(A_{i_1} \cap \dots \cap A_{i_n}) = \sum_{j=1, \dots, n} R(A_{i_j}).$$

A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *consistent* in the sense of R just in case for all $i, j \in I$ and all $V \in \mathcal{V}$:

$$R(X_i^{-1}(V)) > 0 \Leftrightarrow R(X_j^{-1}(V)) > 0$$

Consider a sequence of random variables X_1, \dots, X_n, \dots . The *relative failure* of value v in possibility ω in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$ compares

v to its best competitor v' . It is the difference between the number of v -failures and the minimum of the number of all v' -failures:

$$n - |\{i : X_i(\omega) = v, 1 \leq i \leq n\}| - \min_{v' \in \Upsilon} \{n - |\{i : X_i(\omega) = v', 1 \leq i \leq n\}|\}$$

Let $af(v, X^n(\omega)) = n - afr(v, X^n(\omega))$ be the *absolute failure* of value v in possibility ω in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$. Then the relative failure can be written as:

$$rf(v, X^n(\omega)) = af(v, X^n(\omega)) - \min_{v' \in \Upsilon} \{af(v', X^n(\omega))\}$$

The relative failure of the set of values $V \in \mathcal{V}$ in possibility ω in the finite initial segment $X_1(\omega), \dots, X_n(\omega)$ is defined as follows:

$$rf(V, X^n(\omega)) = \min_{v \in V} \{rf(v, X^n(\omega))\}$$

Relative failures $rf(v, X^n(\omega))$ are completely minimitive ranking functions on any algebra over Υ . This is in some contrast to probability theory, where relative frequencies are merely finitely additive probability measures on some suitable algebra over Υ (van Fraassen 1977).

Relative failures generalize what, in statistics, are known as the *modes* of a sample, viz. the values $v \in \Upsilon$ that have occurred most frequently in the first n repetitions of the experiment X (there may be more than one mode in a sample; in this case statisticians speak of multi-modal variables). The modes of a sample $X^n(\omega)$ are exactly those values $v \in \Upsilon$ that receive relative failure 0. Relative failures generalize this notion by also telling us what the second most frequent outcomes are, and by how less frequent they are. And what the third most frequent outcomes are, and by how less frequent they are. And so on. The thesis that I submit is that modes confirm counterfactuals in much the same way that means and relative frequencies confirm hypotheses about expected values and objective chances.

6 The Obvious Observation

Let me now present the principles that relate the three rank-theoretic notions of subjective grades of disbelief, objective counterfactual distances, and observable absolute and relative failures.

The rank-theoretic analogon of the Principal Principle is the **Royal Rule**. This rule, which is introduced in (Huber 2014), is a normative constraint on a priori grades of disbelief relating subjective grades of disbelief, or conditional beliefs, and objective counterfactual distances, or counterfactuals. The rank-theoretic analogon of the Strong Law of Large Numbers is the **Obvious Observation**. This observation, which will be stated below, is a theorem of the ranking calculus. As a theorem it holds independently of how we interpret the ranking function that figures in it. I will choose to read it as relating objective counterfactual distances and limiting absolute and relative failures (for independent and consistent random variables). We could also formulate the rank-theoretic analoga of the Straight Rule and of the Axiom of Convergence or Reichenbach Axiom. They relate (limiting) relative failures and (limiting) subjective grades of disbelief. However, since they play no role in this paper, I will not formulate them.

The Royal Rule and the Obvious Observation jointly imply that observable absolute failures raise and lower subjective grades of disbelief in hypotheses about objective counterfactual distances. In other words, these two principles jointly imply that modes affect conditional beliefs in counterfactuals. In this sense the Royal Rule and the Obvious Observation tell us what we should believe about what would have been the case.

The **Royal Rule** says that a subject S 's a priori grade of disbelief in some proposition A given that the counterfactual distance to the closest A -worlds equals n – as well as, perhaps, information E that is admissible, but no further information that is inadmissible– should equal n :

$$R_S(A \mid r(A) = n \cap E) = n.$$

Under the assumption that the world w 's theory of deterministic alethic modality or counterfactuality D_w is admissible it follows:

$$r_w(A) = R_S(A \mid D_w).$$

This reformulation says that the counterfactual distance distribution of a world w comes from a subject S 's a priori grading of disbelief by conditionalizing on the world w 's theory of deterministic alethic modality or counterfactuality.⁵

⁵Strictly speaking this notion of admissibility is relative to a context c , just as the notion of admissibility in the Principal Principle is relative to a point of time t . And strictly speaking one needs to assume that the presuppositions of a given context are admissible in this context, and that the world's theory of deterministic alethic modality, or counterfactuality, is admissible in all contexts. However, the context sensitivity of counterfactuals does not play a role for present purposes, and so the presuppositions of all contexts can be assumed to be tautological. It is perhaps worth noting that history up to some time is a context. For details see Huber (2014: sct. 4).

The Obvious Observation derives its name from the fact that it is very easy to prove.⁶ Much like the Strong Law of Large Numbers it considers a sequence of \mathcal{A} - \mathcal{V} -measurable functions X_1, \dots, X_n, \dots , except that now the sequence is assumed to be independent and consistent in the sense of a ranking function rather than independent and identically distributed in the sense of a probability measure.

Theorem 1 (Obvious Observation) *Let R be a completely minimitive ranking function, let X_1, \dots, X_n, \dots be a sequence of \mathcal{A} - \mathcal{V} -measurable functions that is independent and consistent in the sense of R , and let X is an arbitrary one of them.*

There is an $A \in \mathcal{A}$ with $R(\overline{A}) = \infty$ such that for all $\omega \in A$: the relative failure of v in ω in $X_1(\omega), \dots, X_n(\omega)$ will eventually be positive if the rank of v in ω is. That is, if $R(X = x) > 0$, then there is an m such that for all $n \geq m$: $X_n(\omega) \neq x$, i.e. $af(x, X^{n+1}(\omega)) = af(x, X^n(\omega)) + 1$, and so $rf(x, X^n(\omega)) > 0$. For the last claim the range Υ of X has to be finite.

Proof: Suppose X_1, \dots, X_n, \dots is a sequence of independent and consistent \mathcal{A} - \mathcal{V} -measurable functions with $R(X = x) > 0$, where X is an arbitrary one of them. Let D_m^n be the proposition that x has occurred at least m times in X_1, \dots, X_n . Recall that $X = x$ is the proposition $\{\omega \in \Omega : X(\omega) = x\}$. D_m^n is a disjunction of disjuncts of the form $X_1 = x_1 \cap \dots \cap X_n = x_n$ with $x_i = x$ for at least m of the n values x_1, \dots, x_n . Let E be an arbitrary one of those disjuncts whose disjunction is D_m^n . Since X_1, \dots, X_n, \dots is consistent and $R(X = x) > 0$ we have $R(X_i = x) =: r_i > 0$ for all $i = 1, \dots, n$. Let $r = \min\{r_i : i = 1, \dots, n\} \geq 1$. Since X_1, \dots, X_n, \dots is independent we have $R(E) = \sum_{i=1, \dots, n} R(X_i = x_i) \geq m \cdot r \geq m$. Hence $R(D_m^n) = \min\{R(E) : E \text{ is one of the disjuncts of } D_m^n\} \geq m$.

Let $\omega \in \Omega$ and suppose $R(\{\omega\}) < \infty$. x occurs only finitely many times in $X_1(\omega), \dots, X_n(\omega), \dots$. For otherwise for all m there is an n such that $\omega \in D_m^n$, and so $R(\{\omega\}) \geq R(D_m^n) \geq m$, which contradicts the assumption that $R(\{\omega\}) < \infty$. Since this is true for every $\omega \in \Omega$ with $R(\{\omega\}) < \infty$ it follows that there is an $A \in \mathcal{A}$ with $R(\overline{A}) = \infty$ such that this is true for all $\omega \in A$.

This establishes the main claim that there is an m such that for all $n \geq m$: $X_n(\omega) \neq x$, i.e. $af(x, X^{n+1}(\omega)) = af(x, X^n(\omega)) + 1$.

If, in addition, the range Υ of X is finite, then at least one value $v \in \Upsilon$ that is different from x occurs infinitely many times in $X_1(\omega), \dots, X_n(\omega), \dots$. In this case there is an m such that for all $n \geq m$: $rf(x, X^n(\omega)) > 0$. \square

⁶Spohn (2012: ch. 12) proves many much more impressive results that are related to the Obvious Observation. However, the mathematics is not exactly the same, and the interpretation is entirely different. I presently cannot relate his results to mine in an illuminating way.

Even if Υ is not a set of numbers, it is possible to form the modes of X in $X_1(\omega), \dots, X_n(\omega), \{v' \in \Upsilon : rf(v', X^n(\omega)) = 0\}$, as well as the set of values of X that are *typical* or *normal* according to R , or not disbelieved by R ,

$$\{v' \in \Upsilon : R(\{\omega' \in \Omega : X(\omega') = v'\}) = 0\}.$$

Then the Obvious Observation says that the modes eventually track the typical or or normal or non-disbelieved values in all worlds except, perhaps, those in a set of rank infinity.⁷

The Obvious Observation is a theorem of the ranking calculus. Therefore it holds independently of how we interpret the ranking function that figures in it. As already mentioned, I choose to read it as relating (limiting) observable absolute and relative failures and objective counterfactual distances and so arrive at the following result.

For each counterfactual distance space $\mathbb{C} = (\Omega, \mathcal{A}, r)$ whose counterfactual distance distribution r is completely minimitive, each measurable space (Υ, \mathcal{V}) with possibly uncountable Υ , and each sequence of independent and consistent \mathcal{A} - \mathcal{V} -measurable functions X_1, \dots, X_n, \dots there is an $A \in \mathcal{A}$ with $r(\overline{A}) = \infty$ such that for all $\omega \in A$: if x has positive objective rank, then x occurs only finitely many times in $X_1(\omega), \dots, X_n(\omega), \dots$; that is, if $r(X = x) > 0$, then there is an m such that for all $n \geq m$: $X_n(\omega) \neq x$, i.e. $af(x, X^{n+1}(\omega)) = af(x, X^n(\omega)) + 1$.

7 Conditional Independence and Consistency

The sequence of random variables X_1, \dots, X_n, \dots is assumed to be independent and consistent in the sense of the counterfactual distance distribution r . Presumably it will be more difficult to establish that a particular sequence of experiments, or a sequence of repetitions of an experiment type, is independent and consistent in the sense of the counterfactual distance distribution of the actual world than to establish that it is so in the sense of a subject S 's grading of disbelief. Conversely, once these assumptions are established for the objective ranking function of the actual world, the Obvious Observation says something about the counterfactuals that are true in the actual world rather than some subject S 's conditional beliefs.

⁷For the relationship between rank-theoretic typicality or rank-theoretic normality, causation, and counterfactuals see Halpern (2008), Halpern & Hitchcock (2010; 2013), and Huber (2013c).

As in the probabilistic case it is important to note that the \mathcal{A} - \mathcal{V} -measurable functions $X_i : \Omega \rightarrow \Upsilon$ are *singular variables*. They are defined on an exclusive set of possibilities Ω . X_1 may assign the value *tired* to a possible world if I am tired at noon on Monday in this possible world, and \neg *tired* otherwise. X_2 may assign the value *tired* to a possible world if I am tired at noon on Tuesday in this possible world, and \neg *tired* otherwise. X_3 may assign the value *tired* to a possible world if I am tired at noon on Wednesday in this possible world, and \neg *tired* otherwise. And so on. As before, if we have a sequence of such singular variables X_1, \dots, X_n, \dots , we can use this sequence to construct a population $D = \{X_i : i \in \mathbb{N}\}$. We can think of the function $G : D \rightarrow \Upsilon$ as the generic variable that, roughly, assigns the value *tired* or \neg *tired* to an individual at a specified time, me at noon on Monday or me at noon on Tuesday or me at noon on Wednesday, depending on whether or not this individual is tired at noon of the relevant day.

The assumption that the sequence of singular variables X_1, \dots, X_n, \dots is independent means that whether or not X_1 takes on some value x_1 does not affect the objective rank that X_2 takes on some value x_2 . Whether or not I am tired at noon on Monday does not affect the objective rank of whether or not I am tired at noon on Tuesday. The assumption that the sequence of singular variables X_1, \dots, X_n, \dots is consistent means that the objective rank that X_1 takes on the value x is positive just in case the objective rank is positive that X_3 takes on this value x . The objective rank that I am tired at noon on Monday is positive just in case the objective rank is positive that I am tired at noon on Wednesday.

In an intuitive sense rank-theoretic independence and consistency are weaker than probabilistic independence and identical distribution. This sense is intuitive, because we have not assumed a formal connection between counterfactuals and chances. To get a better understanding of the sense in which the rank-theoretic conditions are weaker it will be helpful to consider the notions of *conditional independence* and *conditional consistency*. If we then restrict the discussion to binary variables that take on only two values, effectively: propositional variables, we can reformulate these two assumptions about objective ranking functions as assumptions about counterfactuals. Let us start with the general notions.

A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *independent conditional* on a family of \mathcal{A} - \mathcal{V}^* -measurable functions $(Y_i)_{i \in I}$ ⁸ in the sense of R if and only if for any possible assignment of values $\vec{y} \in \Upsilon^{*I}$ to all variables in $(Y_i)_{i \in I}$, any finite

⁸In general the functions Y_i need not have the same index set I as the first family $(X_i)_{i \in I}$. Nor do they have to have a common range Υ^* and associated algebra \mathcal{V}^* . However, this is the special case we are interested in.

subfamily X_{i_1}, \dots, X_{i_n} , and all $A_{i_1} \in \gamma(X_{i_1}), \dots, A_{i_n} \in \gamma(X_{i_n})$:

$$R(A_{i_1} \cap \dots \cap A_{i_n} \mid \vec{Y} = \vec{y}) = \sum_{j=1, \dots, n} R(A_{i_j} \mid \vec{Y} = \vec{y})$$

A family of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ is *consistent conditional* on a family of \mathcal{A} - \mathcal{V}^* -measurable functions $(Y_i)_{i \in I}$ in the sense of R if and only if for all $i, j \in I$, all $V \in \mathcal{V}$, all $y \in \Upsilon^*$, any possible assignment of values $\vec{y}_{\neq i} \in \Upsilon^{*I \setminus \{i\}}$ to all variables in $(Y_i)_{i \in I}$ except Y_i , and any possible assignment of values $\vec{y}_{\neq j} \in \Upsilon^{*I \setminus \{j\}}$ to all variables in $(Y_i)_{i \in I}$ except Y_j :

$$R(X_i^{-1}(V) \mid Y_i = y \cap \vec{Y}_{\neq i} = \vec{y}_{\neq i}) > 0 \Leftrightarrow R(X_j^{-1}(V) \mid Y_j = y \cap \vec{Y}_{\neq j} = \vec{y}_{\neq j}) > 0$$

Note that the two assignments $\vec{y}_{\neq i}$ and $\vec{y}_{\neq j}$ need not have any value in common, and that i may be identical to j .

Conditional consistency implies, but is not implied by, the assumption that for all $i, j \in I$, all $V \in \mathcal{V}$, and all $y \in \Upsilon^*$:

$$R(X_i^{-1}(V) \mid Y_i = y) > 0 \Leftrightarrow R(X_j^{-1}(V) \mid Y_j = y) > 0$$

Conditional consistency also implies that, if $R(X_i^{-1}(V) \mid Y_i = y) > 0 (= 0)$, then $R(X_i^{-1}(V) \mid Y_i = y \cap \vec{Y}_{\neq i} = \vec{y}_{\neq i}) > 0 (= 0)$, for any possible assignment of values $\vec{y}_{\neq i} \in \Upsilon^{*I \setminus \{i\}}$ to all variables in $(Y_i)_{i \in I}$ except Y_i .

The two families of singular variables $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ are defined on the same domain \mathcal{A} . That is, both families consist of functions that assign values to the possible worlds in Ω . They may differ in the values Υ and Υ^* (respectively the algebras \mathcal{V} and \mathcal{V}^*) they assign to these possible worlds. The second family may say whether or not I have coffee in the morning of various days. The first family may say whether or not I am tired at noon of these days. It is important for conditional consistency that each variable of the first family is uniquely associated with a variable of the second family. This is why the two families of variables have the same index set I . It is important for both conditional consistency and conditional independence (see below) that the condition specifies the value of each variable Y_i as specifically as possible. It is not enough to say of each variable that its value lies in some set of values $V^* \in \mathcal{V}^*$. The condition has to specify *which* value y_i each variable Y_i takes on. It has to be an *atom* of the algebra that is generated by the family $(Y_i)_{i \in I}$ (we can ignore atomless algebras). It is only then, when everything is specified as completely as possible given the expressive

powers of the family $(Y_i)_{i \in I}$, that independence and consistency are supposed to hold.

Now let us consider, in sloppy notation, the special case where both families of random variables consist of two binary variables α, β taking on values $\pm\alpha, \pm\beta$ and γ, δ taking on values $\pm\gamma, \pm\delta$. γ, δ is *independent conditional on α, β* according to the objective ranking function r if and only if

$$r(\pm\gamma \cap \pm\delta \mid \alpha, \beta) = r(\pm\gamma \mid \alpha, \beta) + r(\pm\delta \mid \alpha, \beta).$$

This is to be read as follows: for each of the four possible values for the pair α, β there are exactly four equations that are supposed to hold, one for each of the four possible values for the pair γ, δ .

In still sloppy notation, γ, δ is *consistent conditional on α, β* according to the objective ranking function r if and only if

1. $r(+\gamma \mid +\alpha, \pm\beta) > 0 \Leftrightarrow r(+\delta \mid +\beta, \pm\alpha) > 0$;
2. $r(-\gamma \mid +\alpha, \pm\beta) > 0 \Leftrightarrow r(-\delta \mid +\beta, \pm\alpha) > 0$;
3. $r(+\gamma \mid -\alpha, \pm\beta) > 0 \Leftrightarrow r(+\delta \mid -\beta, \pm\alpha) > 0$; and
4. $r(-\gamma \mid -\alpha, \pm\beta) > 0 \Leftrightarrow r(-\delta \mid -\beta, \pm\alpha) > 0$.

This is to be read as follows: for each of the four possible combinations of values $\pm\alpha$ and $\pm\beta$ there are exactly four equivalences that are supposed to hold, one for each of the four possible combinations of values for γ, δ (that have the same range) and for α, β (that have the same range). Let us focus on the first of those equivalences. It says that, for any arbitrary value of β and any arbitrary value of α , the variable γ takes on its positive value given that α takes on its positive value (and β takes on its specified arbitrary value) just in case δ takes on its positive value given that β takes on its positive value (and α takes on its specified arbitrary value). Similarly for the other three equivalences. This implies, but is not implied by, the assumption

1. $r(+\gamma \mid +\alpha) > 0 \Leftrightarrow r(+\delta \mid +\beta) > 0$;
2. $r(-\gamma \mid +\alpha) > 0 \Leftrightarrow r(-\delta \mid +\beta) > 0$;
3. $r(+\gamma \mid -\alpha) > 0 \Leftrightarrow r(+\delta \mid -\beta) > 0$; and
4. $r(-\gamma \mid -\alpha) > 0 \Leftrightarrow r(-\delta \mid -\beta) > 0$.

Now the first equivalence says: the variable γ takes on its positive value given that α takes on its positive value just in case δ takes on its positive value given that γ takes on its positive value. And similarly for the other three equivalences.

There are eight additional equations for each of the four possible values for the pair α, β , and there are four additional equivalences for each of the four possible combinations of values $\pm\alpha$ and $\pm\beta$, that are supposed to hold, namely those when the values of γ or δ are the tautological or the empty set. However, these equations and equivalences hold trivially. Note that the second definition differs from the first one in that we quantify over the range (Υ and Υ^* respectively \mathcal{V} and \mathcal{V}^*) of the variables rather than their domain (Ω respectively \mathcal{A}), as we do in the first definition.

Think of α, β as specifying the interventions that are involved in performing the experiments γ, δ . Tossing a coin on a given occasion and then checking to see if it lands heads or tails on this occasion. Administering a certain number of mg of ibuprofen to someone in the morning and then checking to see if there is relief of pain shortly afterwards. Making me drink coffee in the morning of some day and then checking to see if I am tired at noon of this day. γ and δ are independent conditional on α and β if, but not only if, the following holds: once it is specified whether I have coffee in the morning of Monday and Tuesday, whether or not I am tired at noon on Tuesday is *counterfactually independent* of whether or not I am tired at noon on Monday. Whether or not I have coffee in the morning of Monday and Tuesday “screens off” any possible counterfactual dependence of me being tired at noon on Tuesday on me being tired at noon on Monday.

That is, once we specify whether I have coffee in the morning of Monday and Tuesday the following four counterfactuals are all false:

1. if I had been tired on Monday, I would have been tired on Tuesday;
2. if I had been tired on Monday, I would not have been tired on Tuesday;
3. if I had not been tired on Monday, I would have been tired on Tuesday; and
4. if I had not been tired on Monday, I would not have been tired on Tuesday.

The corresponding independence assumption in the probabilistic case says that once we specify whether the coin is tossed on the first and second occasion, there is no probabilistic correlation between whether or not it lands heads on the first toss and whether or not it lands heads on the second toss. And once we specify how many mg of ibuprofen are administered to Simone de Beauvoir and Jean-Paul

Sartre, whether or not she feels relief of pain is probabilistically independent of whether or not he feels relief of pain.

The notion of counterfactual independence in the Stalnaker-Lewis approach is defined for propositional variables or binary variables. For this case we can prove that independence in the sense of a ranking function r is implied by, but does not imply, counterfactual independence with r as the counterfactual distance distribution: γ is counterfactually independent of α just in case the following four counterfactuals are all false: $+\alpha \Box \rightarrow +\gamma$, $+\alpha \Box \rightarrow -\gamma$, $-\alpha \Box \rightarrow +\gamma$, $-\alpha \Box \rightarrow -\gamma$.

Let us focus on $+\alpha \Box \rightarrow +\gamma$, as the other three cases are exactly parallel. $+\alpha \Box \rightarrow +\gamma$ is false just in case the r -closest $\llbracket +\alpha \rrbracket$ -worlds are not a subset of the $\llbracket +\gamma \rrbracket$ -worlds. This means $r(\llbracket +\alpha \rrbracket \cap \overline{\llbracket +\gamma \rrbracket}) = r(\llbracket +\alpha \rrbracket)$. The same is true for the other three cases, which implies that $r(\llbracket +\gamma \rrbracket) = r(\llbracket -\gamma \rrbracket) = 0$. So the following four equations hold: $r(\llbracket \pm\alpha \rrbracket \cap \llbracket \pm\gamma \rrbracket) = r(\llbracket \pm\alpha \rrbracket) + r(\llbracket \pm\gamma \rrbracket)$. And this is just the definition of the unconditional independence of the binary variables α and γ according to r . The generalization to conditional independence is straightforward, because any conditional ranking function is a ranking function, and because we condition on one and the same proposition that completely specifies the values of all variables in the second family of variables. Note, though, that *conditional* counterfactual independence is only meaningful on the present account of counterfactuals in terms of ranking functions. Conditional counterfactual independence is not meaningful on an account along the lines of Stalnaker (1968) or Lewis (1973), because there is no such thing as a conditional selection function or a conditional sphere of similarity.

Suppose probabilistic independence implies counterfactual independence. In this case the rank-theoretic independence assumption of the Obvious Observation is strictly weaker than the probabilistic independence assumption of the Strong Law of Large Numbers.⁹ Without this supposition probabilistic independence

⁹Leitgeb's (2012a; b) probabilistic analysis of counterfactuals requires $\Pr(C | A) = 1$ for $A \Box \rightarrow C$ to be true, where \Pr is a Popper-Rényi measure (Popper 1955, Rényi 1955) that is interpreted objectively as time-relative conditional single case chance. Independence in the sense of a Popper-Rényi measure \Pr is a relation between three propositions: A is independent of B conditional on C just in case $\Pr(A \cap B | C) = \Pr(A | C) \cdot \Pr(B | C)$ (Rényi 1970: 103). Counterfactual independence is a relation between two propositions. Therefore the definition of independence in the sense of a Popper-Rényi measure has to be modified so that it becomes a relation between two propositions A and B . It is tempting to say that A is independent of B just in case A is independent of B conditional on the set of all possible worlds, the tautological proposition W . However, on Leitgeb's logic of counterfactuals, just as on ours, $\top \Box \rightarrow \alpha$ is not logically equivalent to α . Therefore this temptation should be resisted. There are other options for modifying independence in the sense of a Popper-Rényi measure so that it becomes a relation between two propositions (see

and rank-theoretic independence are exactly parallel in their unconditional version and in their conditional version. It is not my aim to defend this supposition. However, some philosophers might find the following consideration to make it intuitively plausible. Intuitively, probabilities are modalities that are more fine-grained than counterfactuals. Perhaps one can point to the difference between continuity and discreteness to explain this intuition. If the fine-grained, or high-resolution, modality of probability cannot detect any dependencies, then it would be odd to assume that one can detect dependencies once one has zoomed out to the coarse-grained, or low resolution, modality of counterfactuals. For this reason I have claimed that, in an intuitive sense, probabilistic independence is stronger than counterfactual independence, and hence stronger than independence in the sense of an objective ranking function.

To require that γ, δ are consistent conditional on α, β presupposes that γ is paired with α and that δ is paired with β . Whether or not the coin lands heads or tails on the n th occasion is paired with whether or not the coin is tossed on the n th occasion. Whether or not Simone de Beauvoir feels relief of pain is paired with how many mg of ibuprofen are administered to *her* slightly earlier in the morning. Whether or not I am tired at noon of some day is paired with whether or not I have coffee in the morning of *this* day. In the definition of the general notion this is reflected in the assumption that the two families $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ have the same index set I .

Conditional consistency requires that the same counterfactual relationships obtain between each member of the first family and its associated member from the second family once the values of all other variables in the second family are set to some values in the first case and some values – not necessarily the same – in the other case. That is, conditional consistency requires that, for any arbitrary value for β (say *coffee*) and any arbitrary value for α (say \neg *coffee*), the following counterfactuals for associated pairs α, γ and β, δ have the same truth value: $\pm\alpha (\wedge + \beta) \Box \rightarrow \pm\gamma \Leftrightarrow \pm\beta (\wedge - \alpha) \Box \rightarrow \pm\delta$

(i-a) if I had had coffee in the morning of Monday (and Tuesday), then I would have been tired at noon of Monday JUST IN CASE (i-b) if I had had coffee in the morning of Tuesday (but not Monday), then I would have been tired at noon on

Fitelson & Hájek ms). However, without stipulating which modification one chooses there are no formal relations between counterfactual independence, which relates two propositions, and independence in the sense of a Popper-Rényi measure Pr , which relates three propositions. Another complication may arise from a negative answer to the question whether the set of propositions for which time-relative conditional single case chances are defined is as rich as the set of propositions which can figure in the consequent of a counterfactual.

Tuesday.

(ii-a) if I had not had coffee in the morning of Monday (but on Tuesday), then I would have been tired at noon of Monday JUST IN CASE (i-b) if I had not had coffee in the morning of Tuesday (nor on Monday), then I would have been tired at noon on Tuesday.

(iii-a) if I had had coffee in the morning of Monday (and Tuesday), then I would not have been tired at noon of Monday JUST IN CASE (i-b) if I had had coffee in the morning of Tuesday (but not on Monday), then I would not have been tired at noon on Tuesday.

(iv-a) if I had not had coffee in the morning of Monday (but on Tuesday), then I would not have been tired at noon of Monday JUST IN CASE (iv-b) if I had not had coffee in the morning of Tuesday (nor on Monday), then I would not have been tired at noon on Tuesday.

These four equivalences are supposed to hold for all four possible combinations of an arbitrary value for β and an arbitrary value for α . Suppose they do, i.e. the four counterfactuals $\pm\alpha (\wedge\beta) \square\rightarrow \pm\gamma$ have the same truth values as the four counterfactuals $\pm\beta (\wedge\alpha) \square\rightarrow \pm\delta$ conditional on all four possible combinations of an arbitrary value for β and an arbitrary value for α . Then it follows that the first four counterfactuals have the same truth values as second four counterfactuals: $\pm\alpha \square\rightarrow \pm\gamma \Leftrightarrow \pm\beta \square\rightarrow \pm\delta$

The corresponding identical distribution assumption in the probabilistic case says that the probability that the coin lands heads on the first occasion given that it is tossed on the first occasion (and on the second and third, say) is identical to the probability that the coin lands heads on the second occasion given that it is tossed on the second occasion (but not on the first or third, say). And the probability that Simone feels relief of pain given that she has been administered 300 mg of ibuprofen slightly earlier in the morning (and Jean-Paul and Albert have not been administered any, say) is identical to the probability that Jean-Paul feels relief of pain given that he has been administered 300 mg of ibuprofen slightly earlier in the morning (and so have been Simone and Albert, say).

The argument from above can be repeated to prove that conditional consistency in the sense of a ranking function r is implied by, and (now also) implies, that the various counterfactuals between a given member of the first family and its associated member of the second family have the same truth values once the values of all other variables in the second family are set to some values in the first case and to some, possibly different, values in the other case. If this conditional version of the identity of truth values for counterfactuals holds for all possible combinations of the conditional part, then its familiar unconditional version holds

as well. As before note that *conditional* consistency is only meaningful on the present account of counterfactuals, not on Stalnaker's (1968) or Lewis' (1973).

Suppose identical conditional probabilities imply consistent conditional ranks, i.e. identical conditional truth values for counterfactuals. Then the rank-theoretic consistency assumption of the Obvious Observation is strictly weaker than the identical distribution assumption of the Strong Law of Large Numbers. Note that this supposition is forced upon anyone who defends a probabilistic analysis of counterfactuals, such as Leitgeb (2012a; b).¹⁰ Proponents of such an account are, of course, not forced to rely on the Strong Law of Large Numbers in their account of the testing of counterfactuals. (It will be enough to require that the probability for $X_i^{-1}(V)$ given $Y_i = y$ and some assignment to the rest equals 1 just in case the probability for $X_j^{-1}(V)$ given $Y_j = y$ and some (possibly distinct) assignment to the rest equals 1.)

8 What I Should Believe About What Would Have Been the Case

Let us put to work the Royal Rule and the Obvious Observation that relate the three rank-theoretic notions of subjective grades of disbelief, objective counterfactual distances, and observable absolute and relative failures.

The Royal Rule implies that observable absolute failures raise and lower subjective grades of disbelief in hypotheses about objective counterfactual distances. More specifically, suppose that a subject S 's grading of disbelief R_S satisfies the Royal Rule. Then learning, or conditionalizing on, information about the observable – not relative, but: – absolute failures of sequences of random variables X_1, \dots, X_n, \dots that are independent and consistent in the sense of the objective counterfactual distance distribution r affects the subject S 's grade of disbelief in hypotheses about the objective counterfactual distance distribution r :

$$\begin{aligned}
 R_S(r(X = x) = k \mid af(x, X^n) = j) &= R_S(af(x, X^n) = j \mid r(X = x) = k) + \\
 &\quad + R_S(r(X = x) = k) - R_S(af(x, X^n) = j) \\
 &= (n - j) \cdot k + R_S(r(X = x) = k) - \\
 &\quad - R_S(af(x, X^n) = j) \\
 &\propto^* f^*(k, j, n) + R_S(r(X = x) = k)
 \end{aligned}$$

¹⁰The comparison to Bradley (2012) is complicated by the fact that the latter works with a multidimensional possible worlds semantics.

The first equation results from applications of the axioms of the ranking calculus. The second equation results from an application of the Royal Rule.

Let us get a better understanding of this equation before exploiting it for our purposes. For $k > 0$ the hypothesis $r(X = x) = k$ says that it is strange for x to occur, with k specifying just how strange (the greater k , the more bizarre it is that x occurs). For this case occurrences of x count against the hypothesis $r(X = x) = k$ in the sense that they contribute to an increase in the grade of disbelief in $r(X = x) = k$. They do so the more occurrences of x there are (the greater $n - j$, which is just the absolute frequency of x in the first n repetitions of X) as well as the more bizarre it is for x to occur (the greater k). For $k = 0$ the hypothesis $r(X = x) = k$ says that it is not strange for x to occur. For this case occurrences of x do not count against the hypothesis that $r(X = x) = k$ in the sense that they do not contribute to an increase in the grade of disbelief in $r(X = x) = k$. This is because $(n - j) \cdot k = 0$ for any n and j , if $k = 0$.

Suppose we have a sequence of independent and consistent \mathcal{A} - \mathcal{V} -measurable functions X_1, \dots, X_n, \dots , where X is an arbitrary one of them. Furthermore suppose that, for some $V \in \mathcal{V}$, $r(X = x) = 0$ for all $x \in V$ and $r(X = x') > 0$ for all $x' \in \Upsilon \setminus V$. The Obvious Observation entails that, in all worlds with a finite objective rank, all these values $x' \in \Upsilon \setminus V$ occur at most finitely many times. Let us assume Υ to be finite. Then, in all worlds with a finite objective rank, at least one of the values $x \in V$, say x^+ , occurs infinitely many times. Therefore the grade of disbelief for all hypotheses $r(X = x^+) = k$ with $k > 0$ will grow without bound. In contrast to this the grade of disbelief in $r(X = x^+) = 0$ will never be greater than its prior grade of disbelief $R(r(X = x^+) = 0)$. Therefore, in all worlds w with a finite objective rank, there is an m such that for all $n \geq m$: the hypothesis $r(X = x^+) = 0$ will not be disbelieved by $R(\cdot \mid af(x^+, X^n) = af(x^+, X^n(w)))$. In contrast to this all hypotheses $r(X = x^+) = k$ with $k > 0$ will be disbelieved by $R(\cdot \mid af(x^+, X^n) = af(x^+, X^n(w)))$. This in turn means that in all worlds with a finite objective rank, the hypothesis $r(X = x^+) = 0$, and hence the hypothesis $r(X \in V) = 0$, will be correctly believed after finitely many steps and forever after.

Now suppose we have two families of \mathcal{A} - \mathcal{V} -measurable functions $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ such that the former is independent and consistent conditional on the latter. Furthermore suppose the following counterfactual is true: $Y_i = y \square \rightarrow X_i = x$ or, more generally, $Y_i = y \square \rightarrow X_i \in V$ for some $V \in \mathcal{V}$, where it is important that X_i and Y_i have the same index i . $Y_i = y \square \rightarrow X_i \in V$ is true just in case $r(X_i \notin V \mid Y_i = y) > 0$, which in turn means that $r(X_i = x' \mid Y_i = y) > 0$ for all

$x' \in \Upsilon \setminus V$. Conditional consistency implies $r(X_i = x' \mid Y_i = y_i \cap \vec{Y}_{\neq i} = \vec{y}_{\neq i}) > 0$ for all $x' \in \Upsilon \setminus V$ and for all possible assignment of values $\vec{y}_{\neq i} \in \Upsilon^{*\setminus\{i\}}$ to all variables in $(Y_i)_{i \in I}$ except Y_i .

Instead of counting the occurrences of x , and failures thereof, in X_1, \dots, X_n we now count only those occurrences of x , and failures thereof, in any given X_i where the corresponding variable Y_i with the same index i is set to the relevant value y . For instance, suppose we want to test whether I would have been tired at noon on Friday, if I had had coffee in the morning of Friday. We intervene and make me have coffee in the morning of Monday and Wednesday. We also intervene and make me not have coffee in the morning of Tuesday and Thursday. Then, when considering the data for the counterfactual that I would have been tired at noon on Friday, if I had not had coffee in the morning of Friday, we only consider whether or not I have been tired on Tuesday and Thursday where I did not have coffee in the morning. If we want to test the counterfactual that I would not have been tired at noon on Friday, if I had had coffee in the morning of Friday, then we only consider whether or not I have been tired on Monday and Wednesday where I did have coffee in the morning.

We do the exact same thing in the probabilistic case where we only consider whether the coin has landed heads or tails on those occasions where the coin has been tossed. And we only consider whether Simone and Jean-Paul feel relief of pain on those mornings where they have been administered 300 mg of ibuprofen slightly earlier.

Returning to our argument, we have to assume that the sequence of singular variables Y_1, \dots, Y_n, \dots contains an infinite subsequence $Y_{j_1}, \dots, Y_{j_n}, \dots$ whose values are all set to y . Suppose Y_1, \dots, Y_n, \dots itself contains only occurrences of y . I do not have coffee in the morning of any day and we want to test if I had would have been tired at noon on Friday, if I had not had coffee in the morning of Friday. The coin is tossed on every occasion and we want to test the objective chance of its landing heads. All French existentialists are administered 300mg of ibuprofen in the morning and we want to test if the chance is raised that pain is relieved if 300 mg of ibuprofen are administered slightly earlier.

Given this assumption the Obvious Observation entails that, in all worlds with a finite objective rank conditional on all variables Y_i being set to y , all values $x' \in \Upsilon \setminus V$ occur at most finitely many times. Again let us assume Υ to be finite. Then, in all worlds with a finite objective rank conditional on all variables Y_i being set to y , at least one value $x \in V$, say x^+ , occurs infinitely many times. Therefore the grade of disbelief for all hypotheses $r(X_i = x^+ \mid Y_i = y \cap \vec{Y}_{\neq i} = \vec{y}) =$

k with $k > 0$ will grow without bound. In contrast to this the grade of disbelief in $r(X_i = x^+ | Y_i = y \cap Y_{\neq i}^{\vec{y}} = \vec{y}) = 0$ will never be greater than its prior grade of disbelief $R(r(X_i = x^+ | Y_i = y \cap Y_{\neq i}^{\vec{y}} = \vec{y}) = 0)$.

Conditional consistency implies that this is also true if $Y_{\neq i}^{\vec{y}} = \vec{y}$ is replaced by any other possible assignment of values $y_{\neq i}^{\vec{y}} \in \Upsilon^{\setminus \{i\}}$ to all variables in $(Y_i)_{i \in I}$ except Y_i . In addition conditional consistency implies that this is also true if we ignore all variables in $(Y_i)_{i \in I}$ except Y_i and just consider $r(X_i = x^+ | Y_i = y)$.

Hence in all worlds w with a finite objective rank there is an m such that for all $n \geq m$: the hypothesis $r(X_i = x^+ | Y_i = y) = 0$ will not be disbelieved by the posterior disbelief function $R(\cdot | af(x^+, X^n | y) = af(x^+, X^n(w | y)))$. In contrast to this all hypotheses $r(X_i = x^+ | Y_i = y) = k$ with $k > 0$ will be disbelieved by the posterior disbelief function $R(\cdot | af(x^+, X^n | y) = af(x^+, X^n(w | y)))$. (The condition y in $af(x^+, X^n | y)$ and $af(x^+, X^n(w | y))$ is to remind us to count only those occurrences of x^+ , and failures thereof, in any given X_i where the corresponding variable Y_i is set to y .) This in turn means that in all worlds with a finite objective rank, the hypothesis $r(X_i = x^+ | Y_i = y) = 0$, and hence the hypothesis $r(X_i \in V | Y_i = y) = 0$, will be correctly believed after finitely many steps and forever after. And this means that in all worlds with a finite objective rank, the true counterfactual $Y_i = y \sqcap \rightarrow X_i \in V$ will, correctly, *not* be disbelieved after finitely many steps and forever after.

Since Υ is finite there are only finitely many counterfactuals of the form $Y_i = y \sqcap \rightarrow X_i \in V$ for $V \subseteq \Upsilon$. This implies that there is a point from which on only false counterfactuals will be disbelieved. The converse is not true, though. If the agent is too cautious, then there may be false counterfactuals that she may never start to disbelieve. Additional principles need to be obeyed by the subject to eventually disbelieve every false counterfactual.

The reader will already have noticed that the worlds with a finite objective rank are exactly those worlds that are accessible in the sense of the modal logic that is contained in any logic of counterfactuals via the following principle: $\Box \alpha \leftrightarrow (\neg \alpha \sqcap \rightarrow \alpha)$. This allows us to reformulate our main result two more times. In all worlds that are accessible, any true counterfactual will, correctly, not be disbelieved after finitely many steps and forever after. Put differently, it is necessary that any true counterfactual will, correctly, not be disbelieved after finitely many steps and forever after. This holds, provided the agent obeys the Royal Rule and conditionalizes on the observable absolute failures of a sequence of experiments that is independent and consistent conditional on the interventions involved in carrying out the experiments.¹¹

9 Comparisons

In concluding this paper I want to briefly compare the account of the testing of counterfactuals presented here to two alternatives.

The first alternative arises from the similarity approach to counterfactuals. There are several versions of this approach, but they all agree that the logic of counterfactuals does not merely validate the basic conditional logic **V**, as, among others, Leitgeb (2012a; b) and Huber (2014) have it, but at least the logic **VW** that arises from **V** by adding weak centering: $(\alpha \Box \rightarrow \gamma) \supset (\alpha \supset \gamma)$. For instance, both Stalnaker’s logic **VCS** and Lewis’ “official logic” **VC** are strictly stronger than **VW**. If the logic of counterfactuals includes weak centering *in the present framework*, then the testing of counterfactuals is much easier: consistency implies that a single experiment is enough to find out that a false counterfactual is false. If the logic of counterfactuals includes strong centering, $(\alpha \wedge \gamma) \supset (\alpha \Box \rightarrow \gamma)$, which is also validated by both **VCS** and **VC**, consistency implies that a single experiment is enough to find out that a true counterfactual is true.

This is true if these principles are included in the logic of counterfactuals in the present framework. However, what distinguishes the present framework from the similarity approach are (i) the doxastic state representing conditional beliefs, the subject *S*’s grading of disbelief $R(H | E)$, and (ii) the normative principle that relates this doxastic state of conditional form to counterfactuals, the Royal Rule. Proponents of the similarity approach are, of course, free to come up with such a principle, as well as a theorem such as the Obvious Observation, that jointly allow them to state under which conditions counterfactuals can be empirically tested or confirmed. However, to the best of my knowledge no such package of principle plus theorem has ever been proposed. Moreover, I am skeptical that it is even possible to formulate a principle analogous to the Royal Rule, as the latter essentially relies on the conditional nature of rank-theoretic grades of (dis)belief, and neither selection functions nor systems of spheres have this conditional nature.

The other alternative is what I have called probabilification by reduction in section 4: reduce a subject *S*’s conditional belief $Bel_S(H | E)$ in hypothesis *H* given various data *E* to the subject *S*’s credence $Cr_S(H | E)$ in this hypothesis *H* given those data *E* along the lines of Leitgeb (2013); reduce counterfactual dependence of certain propositions *C* on certain propositions *A*, $\pm A \Box \rightarrow \pm C$, to the sets of possible worlds *w* where the objective chances of these propositions

¹¹Modal logicians might be interested in the fact that this provides an example of a sentence α such that: $\vdash \Box \alpha$ and $\not\vdash \alpha$.

$\pm C$ given those propositions $\pm A$ are 1, $ch_w(\pm C | \pm A) = 1$, along the lines of Leitgeb (2012a; b) – or along the multidimensional lines of Bradley (2012); and reduce what one observes to happen, and to fail to happen, to a report of those observables in the form of relative frequencies.

This probabilistic alternative can identify the relevant epistemic state of conditional form as the subject S 's credence function $Cr_S(H | E)$. Counterfactuals are given a probabilistic analysis in terms of objective chances. Therefore the normative principle linking this doxastic state of conditional form to counterfactuals is the familiar Principal Principle. The relevant theorem is the Strong Law of Large Numbers, or some theorem with weaker assumptions as suggested in section 5. Observable relative frequencies can then be used to empirically test or confirm counterfactuals as explained in section 3, provided that the story told in section 3 can be carried over from classical probabilities to Popper-Rényi measures. (As indicated in footnote 8, this is not straightforward, but necessary if one adopts Leitgeb's (2012a; b) probabilistic analysis of counterfactuals.)

As mentioned there is no difference between the logic of counterfactuals of this probabilistic alternative and the logic of the present rank-theoretic approach: the basic conditional logic \mathbf{V} is sound and complete with respect to both of them. So this probabilistic alternative also allows for the testing of counterfactuals, just as the present approach does. It seems to be superior to the present approach by being more parsimonious, both epistemologically and ontologically. This is so, because the present approach is not intended as a replacement of probabilistic metaphysics and epistemology, but as a complementing addendum.

There are, however, some subtle differences, and I would like to conclude by briefly pointing these out. The first point is that rank-theoretic independence is strictly weaker than probabilistic independence, given a probabilistic analysis of counterfactuals that requires $\Pr(A \cap C) = \Pr(A)$ for $A \Box \rightarrow C$ to be true. For, under this assumption, probabilistic independence implies, but is not implied by, counterfactual dependence. As explained in section 7, the latter implies rank-theoretic independence. (I am ignoring propositions with extreme probabilities. These are probabilistically independent of all propositions, including themselves. However, presumably, they are counterfactually dependent on all propositions, including themselves, on a probabilistic semantics for counterfactuals.)

The second point is that rank-theoretic consistency is strictly weaker than the identical distribution assumption in the probabilistic case. (It will be possible to weaken this assumption, but I do not think that one can weaken the probabilistic independence assumption). The reason is that identically distributed probabilities imply, but are not implied by, identical truth values of counterfactuals. The

latter in turn imply rank-theoretic consistency as explained in section 7. The upshot is that it is easier to test counterfactuals on the present approach than on the probabilistic alternative that probabilifies by reduction.

Finally suppose we have two sequences of variables $(X_i)_{i \in I}$ and $(Y_i)_{i \in I}$ such that the former is independent and consistent conditional on the latter, as well as independent and identically distributed conditional on the latter. Furthermore, suppose $Y_i = y \square \rightarrow X \in V$ is true, Υ is finite, and all variables Y_i are set to y . On the present account it follows that, necessarily, any $x' \in \Upsilon \setminus V$ occurs only finitely many times. On the probabilistic alternative it follows that, with maximal objective chance, the relative frequency of all $x' \in \Upsilon \setminus V$ converges to 0. But that is compatible with all $x' \in \Upsilon \setminus V$ occurring infinitely many times, even when we are in a world with a finite objective rank and positive objective chance. In this sense counterfactuals are stronger on the present account than on the probabilistic alternative, even though both are characterized by the basic conditional logic **V**.

This means that the present approach comfortably sits between the first alternative that arises from the similarity approach and the second alternative of probabilification by reduction. The first alternative does not allow for any “counter-instances”, as we may call occurrences of some $x' \in \Upsilon \setminus V$, at all. The probabilistic alternative allows for infinitely many counter-instances. The present rank-theoretic approach allows for counter-instances, but only for finitely many.

Acknowledgments

I am very grateful to an anonymous referee, Alan Hájek, Christopher R. Hitchcock, Hannes Leitgeb, Timothy Williamson, and, especially, to Wolfgang Spohn for many most helpful comments and suggestions on several earlier versions of this paper.

References

- [1] Adams, Ernest W. (1998), *A Primer of Probability Logic*. Stanford: CLSI Publications.
- [2] Alchourrón, Carlos E. & Gärdenfors, Peter & Makinson, David (1985), On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic* **50**, 510-530.

- [3] Baumgartner, Michael & Glynn, Luke (2013), Introduction to Special Issue on ‘Actual Causation’. *Erkenntnis* **78**, 1-8.
- [4] Boutilier, Craig (1996), Iterated Revision and Minimal Change of Belief. *Journal of Philosophical Logic* **25**, 263-305.
- [5] Bradley, Richard (2012), Multidimensional possible-world semantics for conditionals. *Philosophical Review* **121**, 539-571.
- [6] Brössel, Peter, & Eder, Anna-Maria, & Huber, Franz (2013), Evidential Support and Instrumental Rationality. *Philosophy and Phenomenological Research* **87**, 279-300.
- [7] Briggs, Rachael (2009), The Big Bad Bug Bites Anti-Realists About Chance. *Synthese* **167**, 81-92.
- [8] Briggs, Rachael (2012), Interventionist Counterfactuals. *Philosophical Studies* **160**, 139-166.
- [9] Carnap, Rudolf (1980), A Basic System of Inductive Logic, Part 2. In R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. Vol. II. Berkeley: University of Berkeley Press, 7-155.
- [10] Collins, John & Hall, Ned & Paul, L.A. (eds., 2004), *Causation and Counterfactuals*. Cambridge, MA: MIT Press.
- [11] Darwiche, Adnan & Pearl, Judea (1997), On the Logic of Iterated Belief Revision. *Artificial Intelligence* **89**, 1-29.
- [12] Earman, John (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- [13] Edgington, Dorothy (2008), Counterfactuals. *Proceedings of the Aristotelian Society* **108**, 1-21.
- [14] Fitelson, Branden & Hájek, Alan (ms), Declarations of Independence.
- [15] Galles, David & Pearl, Judea (1998), An Axiomatic Characterization of Causal Counterfactuals. *Foundations of Science* **1**, 151-182.
- [16] Halpern, Joseph Y. (2008), Defaults and Normality in Causal Structures. *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR 2008)*, 198-208.

- [17] Halpern, Joseph Y. (2013), From Causal Models to Counterfactual Structures. *The Review of Symbolic Logic* **6**, 305-322.
- [18] Halpern, Joseph Y. & Hitchcock, Christopher R. (2010), Actual Causation and the Art of Modelling. In R. Dechter & H. Geffner & J. Halpern (eds.), *Heuristics, Probability, and Causality*. London: College Publications, 383-406.
- [19] Halpern, Joseph Y. & Hitchcock, Christopher R. (2013), Compact Representations of Extended Causal Models. *Cognitive Science* **37**, 986-1010.
- [20] Hild, Matthias & Spohn, Wolfgang (2008), The Measurement of Ranks and the Laws of Iterated Contraction. *Artificial Intelligence* **172**, 1195-1218.
- [21] Hintikka, Jaakko (1961), *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca, NY: Cornell University Press.
- [22] Huber, Franz (2007), The Consistency Argument for Ranking Functions. *Studia Logica* **86**, 299-329.
- [23] Huber, Franz (2013a), Belief Revision I: The AGM Theory. *Philosophy Compass* **8**, 604-612.
- [24] Huber, Franz (2013b), Belief Revision II: Ranking Theory. *Philosophy Compass* **8**, 613-621.
- [25] Huber, Franz (2013c), Structural Equations and Beyond. *The Review of Symbolic Logic* **6**, 709-732.
- [26] Huber, Franz (2014), New Foundations for Counterfactuals. *Synthese*.
- [27] Jeffrey, Richard C. (1970), Dracula Meets Wolfman: Acceptance vs. Partial Belief. In M. Swain (ed.), *Induction, Acceptance, and Rational Belief*. Dordrecht: D. Reidel, 157-185.
- [28] Leitgeb, Hannes (2012a), A Probabilistic Semantics for Counterfactuals. Part A. *Review of Symbolic Logic* **5**, 26-84.
- [29] Leitgeb, Hannes (2012a), A Probabilistic Semantics for Counterfactuals. Part B. *Review of Symbolic Logic* **5**, 85-121.

- [30] Leitgeb, Hannes (2013), Reducing Belief *Simpliciter* to Degrees of Belief. *Annals of Pure and Applied Logic* **164**, 1338-1389.
- [31] Lewis, David K. (1973), *Counterfactuals*. Cambridge, MA: Harvard University Press.
- [32] Lewis, David K. (1979), Counterfactual Dependence and Time's Arrow. *Noûs* **13**, 455-476.
- [33] Lewis, David K. (1980), A Subjectivist's Guide to Objective Chance. In R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. Vol. II. Berkeley: University of Berkeley Press, 263-293.
- [34] Lin, Hanti & Kelly, Kevin T. (2012), Propositional Reasoning that Tracks Probabilistic Reasoning. *Journal of Philosophical Logic* **41**, 957-981.
- [35] Mumford, Stephen (1998), *Dispositions*. Oxford: Oxford University Press.
- [36] Nozick, Robert (1981), *Philosophical Explanations*. Oxford: Oxford University Press.
- [37] Paul, L.A. (2000), Aspect Causation. *Journal of Philosophy* **XCVII**, 235-256.
- [38] Pearl, Judea (2009), *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press.
- [39] Percival, Philip (2002), Epistemic Consequentialism. *Proceedings of the Aristotelian Society*, Supplementary Volume **76**, 121-151.
- [40] Popper, Karl R. (1955), Two autonomous axiom systems for the calculus of probabilities. *British Journal for the Philosophy of Science* **6**, 51-57.
- [41] Reichenbach, Hans (1939), *Experience and Prediction. An Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press.
- [42] Rényi, Alfred (1955), On a New Axiomatic System for Probability. *Acta Mathematica Academiae Scientiarum Hungaricae* **6**, 285-335.
- [43] Rényi, Alfred (1970), *Foundations of Probability*. San Francisco: Holden-Day.

- [44] Spirtes, Peter & Glymour, Clark & Scheines, Richard (2000), *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press.
- [45] Spohn, Wolfgang (1988), Ordinal Conditional Functions: A Dynamic Theory of Epistemic States. In W.L. Harper & B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II*. Dordrecht: Kluwer, 105-134.
- [46] Spohn, Wolfgang (2010), Chance and Necessity: From Humean Supervenience to Humean Projection. In E. Eells & J. Fetzer (eds.), *The Place of Probability in Science. Boston Studies in the Philosophy of Science 284*. Dordrecht: Springer, 101-131.
- [47] Spohn, Wolfgang (2012), *The Laws of Belief. Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.
- [48] Spohn, Wolfgang (2013), A Ranking-Theoretic Approach to Conditionals. *Cognitive Science 37*, 1074-1106.
- [49] Stalnaker, Robert. C. (1968), A Theory of Conditionals. In N. Rescher (ed.), *Studies in Logical Theory. American Philosophical Quarterly Monograph Series 2*, Oxford: Blackwell, 98-112.
- [50] Stalnaker, Robert C. (2002), Epistemic Consequentialism. *Proceedings of the Aristotelian Society, Supplementary Volume 76*, 153-168.
- [51] van Fraassen (1977), Relative Frequencies. *Synthese 34*, 133-166.
- [52] Woodward, James F. (2003), *Making Things Happen*. New York: Oxford University Press.
- [53] Zhang, Jiji (2013), A Lewisian Logic of Causal Counterfactuals. *Minds and Machines 23*, 77-93.
- [54] Zhang, Jiji & Lam, Wai-Yin & De Clercq, Rafael (2013), A Peculiarity in Pearl's Logic of Interventionist Counterfactuals. *Journal of Philosophical Logic 42*, 783-794.