Counterfactual Contamination[1]

Simon Goldstein and John Hawthorne

Keywords: Knowledge, Safety, Counterfactuals

Abstract

Many defend the thesis that when someone knows p, they couldn't easily have been wrong about p. But the notion of easy possibility in play is relatively under-theorized. One structural idea in the literature, the principle of Counterfactual Closure (CC), connects easy possibility with counterfactuals: if it easily could have happened that p, and if p were the case, then q would be the case, it follows that it easily could have happened that q. We first argue that while CC is false, there is a true restriction of it to cases involving counterfactual dependence on a coin flip. The failure of CC falsifies a model where the easy possibilities are counterfactually similar to actuality. Next, we show that extant normality models, where the easy possibilities are the sufficiently normal ones, are incompatible with the restricted CC thesis involving coin flips. Next, we develop a new kind of normality theory that can accommodate the restricted version of CC. This new theory introduces a principle of Counterfactual Contamination, which says roughly that any world is fairly abnormal if at that world very abnormal events counterfactually depend on a coin flip. Finally, we explain why coin flips and other related events have a special status. A central take home lesson is that the correct principle in the vicinity of Safety is importantly normality-theoretic rather than (as it is usually conceived) similarity-theoretic.

1. Introduction

Many defend the thesis that knowledge requires safety from error. If S knows p, then S couldn't easily have believed p falsely.[2] Or, in other terms:

---

[1] This is a preprint of an article whose final and definitive form will be published in the Australasian Journal of Philosophy 2021. The Australasian Journal of Philosophy is available online at: http://www.tandf.co.uk/journals/.

[2] For defenses of Safety, see among others [Sosa 1999; Williamson 2000; Pritchard 2005; Manley 2007; Williamson 2009; Pritchard 2012; Goldstein and Waxman 2020]. Safety principles are standardly relativized to methods. An agent knows p using method m only if they couldn't easily have falsely believed p using method m. We'll mostly suppress method-relativity in what follows, because our arguments could easily be adapted to this setting. Finally, in ordinary language the expression 'could easily' sometimes suggests a requirement of sufficiently high probability. We ignore this effect throughout, following the literature in interpreting 'could easily' as helping to specify a range of possible worlds throughout which an agent must believe reliably.

Safety S knows p only if there is no easy possibility where S believes p falsely.[3]

Goodman [2015] relates easy possibility to counterfactuals, endorsing the following bridge principle:

Counterfactual Closure If there is an easy possibility where p, and if p had been the case, then q would have been the case, then there is an easy possibility where q.[4]

Goodman shows that Safety and Counterfactual Closure make trouble for Lewisian theories of counterfactuals. According to Lewis [1979], there are widespread 'counterfactual miracles', so that even if the world had been just slightly different, the laws of nature would have been different, because there would have been a small miracle that made the world different. Given Safety and Counterfactual Closure, counterfactual miracles leads to skepticism about the laws of nature.

Here's the problem. First, it is possible for scientific inquiry to arrive at knowledge of the laws of nature in a deterministic world. Imagine we lived in a world where some deterministic physical theory was correct, say Newtonian mechanics. Then scientists could have already learned Newtonian mechanics was correct. Scientists would have run the same experiments testing

---

[3]Throughout, the notion of easy possibility is agent and time relative. Our question is which possibilities are easy for the agent at the time of their knowing. Other work uses the term 'nearby possibility' for these worlds. We use the term 'easy possibility' for continuity with [Goodman 2015], and because to our ears 'nearby possibilities' suggests a similarity based theory of easy possibility, which we'll criticize later.

[4]The reader might be tempted to accept analogous principles stated in terms of ignorance rather than easy possibility. We caution against such temptation. Here are a few tries we have encountered. First, if you're ignorant whether p and if p were the case, q would be the case, then you're ignorant whether q. This founders on the fact that every p counterfactually implies every tautology. Second: if you're in no position to know whether p and you believe q and if p were the case, you would falsely believe q (by the same method), then you don't know q. This fails given the vacuous truth of counterfactuals with metaphysical impossible antecedents. Suppose Goldbach's Conjecture is true, and hence necessarily true. Now take any claim q Charlie apparently knows. Charlie isn't in a position to know whether Goldbach's Conjecture is false. But if it were false, Charlie would have falsely believed q (by the same method). So this principle absurdly implies Charlie doesn't know q after all. Here's a third principle: if you don't know whether p and p is metaphysically possible and you believe q and if p were the case, you would falsely believe q (by the same method), then you don't know q. This principle is quite skeptical. Charlie doesn't know whether she recently inhaled polonium, because he knows little about chemistry. Charlie believes that he'll be alive tomorrow. As a matter of fact, polonium is deadly. If Charlie had inhaled polonium, he would die today. This principle implies skeptically that Charlie doesn't know he'll be alive tomorrow. Counterfactual Closure has no such consequence, because there may be no easy possibility in which Charlie recently inhaled polonium. On this proposal, Safety is necessary but not sufficient for knowledge, and so Charlie can be ignorant of whether he inhaled polonium even though there is no easy possibility where he falsely believes this.

quantum mechanics and relativity. But the experiments would have gone differently, confirming Newtonian mechanics.

Lewis's theory of counterfactuals makes heavy use of counterfactual miracles in deterministic worlds. If the recent past had been slightly different, the initial conditions would have been the same. But there would have been a small violation of the laws of nature soon before the time at which the past diverged from actuality. Now imagine a small event such as my not blinking at t, which occurred the day after scientists confirmed Newtonian mechanics. Lewis introduces a counterfactual miracle:

(1) If I had blinked at t, Newtonian mechanics would have been false.

The scientists' method of confirming Newtonian mechanics have nothing to do with my blinking at t. So whether or not I blinked at t is independent of whether or not the scientists would believe in Newtonian mechanics. That is:

(2) If I had blinked at t, the scientists would have falsely believed in Newtonian mechanics.

But (2), Counterfactual Closure, and Safety lead to skepticism about the laws. There is an easy possibility where I blink at t. After all, the scientists have no idea whether I blinked. So Counterfactual Closure implies there is an easy possibility where the scientists believed Newtonian mechanics falsely. So Safety implies that the scientists do not know that Newtonian mechanics is correct. Generalizing, counterfactual miracles lead to skepticism about the laws.

Bennett [1984] and Dorr [2016] develop approaches to counterfactuals that do without such miracles. On these theories, small changes in the past counterfactually imply small changes to the initial conditions of the universe. If I had blinked at t, the initial conditions would have been just different enough to allow my blinking at t. The laws would have remained the same, and the macroscopic features of the world would have been roughly the same throughout time, modulo my blinking. On the basis of the argument above, Goodman [2015] suggests this alternative approach to counterfactuals is superior to a miracle based analysis.

Counterfactual Closure is validated by a model of knowledge that understands knowledge in terms of counterfactual similarity (see for example [Bacon 2014]). In this model, a world is easily possible just in case it is sufficiently similar to actuality. The counterfactual p > q is true just in case the most similar p worlds are q worlds.[5] This model validates Counterfactual Closure. Suppose there is an easy possibility where p. Then there is a sufficiently similar world where p. Suppose p > q. Since there is a sufficiently similar world where p, the most similar p worlds are themselves sufficiently similar. Since p > q is true, it follows that the most similar p worlds are q

[5]We follow Stalnaker [1968] in making the limit assumption, but follow Lewis [1973] in allowing ties for similarity. We could make the same points without the limit assumption, so long as the similarity ordering used to judge counterfactual closeness is the same as that used to judge easy possibility.

worlds. So these worlds are themselves sufficiently similar. So there must be an easy possibility where q.

Below, we first argue that CC fails. This falsifies the model above, where the easy possibilities are those counterfactually similar to actuality. We next show that normality models of knowledge correctly predict that Counterfactual Closure fails. In these models, the easy possibilities are the sufficiently normal worlds. We then offer a true restriction of Counterfactual Closure. This principle says p is an easy possibility when p counterfactually depends on a coin flip. We show that restricted Counterfactual Closure is invalidated by extant normality theories. Next, we develop a new kind of normality theory that can accommodate the restricted version of CC. This new theory introduces a principle of Counterfactual Contamination, which says roughly that any world is fairly abnormal if at that world very abnormal events counterfactually depend on a coin flip. Finally, we explain why coin flips and other related events have a special status. A central take home lesson is that the correct principle in the vicinity of Safety is importantly normality-theoretic rather than (as it is usually conceived) similarity-theoretic.

2. Against Counterfactual Closure

Counterfactual Closure is false.

At noon, a fair coin is flipped and Betty bets a hundred dollars that it will land heads. Betty won't learn the result until 1 o'clock. So at noon, Betty doesn't know whether she will be sad when she sees the result. This means that there's an easy possibility in which Betty is sad when she sees the result. As a matter of fact, the coin landed heads. So if Betty were sad when she saw the result, she'd be sad about winning the bet. But at noon, Betty knows that she won't be sad about winning the bet. So there is no easy possibility where she is sad about winning the bet.[6]

There is a room connected to a cage by a passage. A tiger prowls back and forth between the room and the cage. There is a trap door above the room. If Alex opens the trap door, the

[6] Strictly speaking, this last inference does not follow without further background assumptions. Safety implies that if Betty knows she won’t be sad about winning the bet, then there is no easy possibility where both (i) Betty is sad about winning the bet; and (ii) Betty believes that she won't be sad about winning the bet. But this allows that there could be an easy possibility where Betty is sad about losing the bet, as long as she doesn't believe he won’t be sad about losing the bet in that scenario. First, in this case it is plausible that Betty could not easily have failed to believe that she won’t be sad about winning the bet. After all, Betty knows that she likes to win bets, and this knowledge is modally robust. Since Betty believes she won’t be sad about winning the bet throughout the easy possibilities, an easy possibility in which she is sad about winning the bet is an easy possibility in which she falsely believes she won’t be sad about winning the bet. Second, in this case the following counterfactual is also true: if Betty were going to be sad about winning the bet, she would have falsely believed that wasn’t going to be sad about winning the bet. Another application of Counterfactual Closure thus implies that if there is an easy possibility where Betty is sad about winning the bet, then there is also an easy possibility where Alex falsely believes he won't be sad about winning the bet. Similar remarks apply to the next example.

passage seals, locking the tiger in its current location. Alex plans to peek into the room and then spend the night just in case the tiger isn't there. In fact, the trap door is slightly ajar and so the passage is closed. The tiger is trapped in the room. Right before Alex peeks, there is a paradigmatically easy possibility that the tiger is in its cage. Alex knows he'll spend the night just in case the tiger is in its cage. So there's an easy possibility that he'll spend the night in the room. But since there is actually a tiger in the room, Alex would have been eaten by the tiger if he spent the night. But Alex knows he won't be eaten by a tiger. So there is no easy possibility where he's eaten by a tiger.

We rely on the counterfactual that if Betty were sad when she saw the result, she'd be sad about winning the bet. We also rely on the counterfactual that if Alex were to spend the night in the room, he would be eaten by the tiger. These counterfactuals are common ground between various standard approaches to counterfactuals. These counterfactuals don't decide whether counterfactual interventions require miracles versus tweaked initial conditions. They reflect our inclination to hold the outcome of a past coin toss when considering changes to later facts regarding someone's emotions. Likewise, the second counterfactual reflects our inclination to hold fixed where the tiger is when considering changes to where Alex is.[7]

The model of easy possibility in terms of similarity cannot explain these cases. This model must rank the relative similarity of worlds where Betty is sad about losing and worlds where Betty is sad about winning. To match our judgments about easy possibility, worlds where Betty is sad about losing must be significantly more similar than worlds where she is sad about winning. But then it is false that if she had been sad when she saw the result, she would have been sad about winning.[8] [9]

3. Few Easy Possibilities

One response to our counterexamples denies various easy possibilities. Perhaps there is no easy possibility where Alex stays in the room. After all, there will be a tiger in the room. Perhaps

---

[7] You might think the common thread here is a no-backtracking constraint on counterfactual evaluation. While that captures Lewis's theory of counterfactuals, it doesn't capture Dorr's approach, where we do allow certain fine details about the past to counterfactually depend on the future. But even Dorr's approach is crafted to preclude the outcome of past coin-tosses depending on future emotion, and to preclude the room in which the tiger resides depending on Alex' future behavior. Dorr rightly takes it as a datum, for example, that if we were to act strangely tomorrow, the macroscopic lay of the land today would have been the same.

[8] Although we reject Counterfactual Closure, we have not refuted a weakening of Counterfactual Closure. According to this weakening, Counterfactual Closure applies when the relevant agent knows the counterfactual if p were the case, then q would be the case. Our counterexamples to Counterfactual Closure are not counterexamples to this weakened principle. In each case, the agent is ignorant of the relevant counterfactual.

[9] Dorr [2016] offers other arguments against counterfactual miracles, which appeal to the connection between counterfactuals and emotions. Our counterexamples to Counterfactual Closure do not falsify those arguments.

there is no easy possibility where Betty is sad when she sees the result. After all, she won the bet.

This response is consistent if Safety is merely a necessary condition for knowing. In that case, an agent can be ignorant of p even though she couldn't easily have believed p falsely. This response agrees that Alex doesn't know whether he will enter the room. The response agrees that Betty doesn't know whether she will be sad about the result of her bet. But the response claims there is no easy possibility where Alex doesn't enter the room, and no easy possibility where Betty is sad about the result.

Here is a plausible sufficient condition for easy possibility. If an agent knows a fair coin is about to be flipped, and p depends on the outcome of the flip, then there is an easy possibility for that agent where p. Say a fair coin flip is 'active' when the relevant agent knows it is about to take place. Then the outcomes of active coin flips are easy possibilities; and other events that depend on them are easy possibilities also.

We can adapt our counterexamples so that the relevant outcome depends on an active coin flip. Imagine the tiger's location is controlled not by a locking passage, but by a coin flip. Imagine Alex knows this. At noon, Alex doesn't know whether he'll spend the night in the room. A coin flip at 12:30 will determine whether the tiger is in the room: the tiger will be in the room iff the coin lands heads. So there is an easy possibility where Alex stays in the room. As a matter of fact, that coin will land heads. So as a matter of fact, Alex won't spend the night in the room, and he would have been eaten by the tiger if he had. Nonetheless, Alex knows at noon that he won't be eaten by a tiger. After all, if there is a tiger in the room he will notice when he peeks, and he will choose not to stay in the room.[10]

An epistemic heretic might deny that active coin flips make for easy possibilities. But there is a more general point here. The denial of too many easy possibilities trivializes Safety. Safety says that knowledge requires believing truly across the easy possibilities. If there are very few easy possibilities, then Safety explains very few cases of ignorance. Imagine there's a .99 biased heads coin flip that just occurred, and landed heads. You can't see the coin. Plausibly, you don't know that the coin landed heads. But you have a justified high credence that the coin landed heads. To explain your ignorance, it is natural to appeal to Safety. You easily could have believed falsely that the coin landed heads, because the coin easily could have landed tails.

[10] A sophisticated opponent might deny that in each case, the relevant counterfactual is true at noon. On this proposal, the possibility of the coin landing tails blocks the truth at noon of the counterfactual that if Alex had entered the room, he would have been eaten by the tiger. A natural way to develop this idea is to say that the counterfactual changes truth value over time, so that it is true only after the coin lands. To handle this opponent, we can strengthen our sufficient condition so that if a fair coin was flipped a moment ago and the outcome hasn't been viewed, there is an easy possibility now that the coin landed heads and an easy possibility that it landed tails. Then we can imagine that Alex and Billy's conditions were determined by a coin that just flipped, but where they can't see the outcome of the flip yet.

This case is best explained by easy possibility, and just this kind of easy possibility falsifies Counterfactual Closure.

4. Normality

Our counterexamples rule out the similarity model of easy possibility, which implies Counterfactual Closure. We'll now see that models of easy possibility in terms of normality predict the failure of Counterfactual Closure.

A variety of recent work models knowledge via normality.[11] Here is Beddor and Pavese [2019: 12-13]: 'what determines which worlds count as relevantly close? Here is one suggestion. When we assess a task such as shooting hoops, we implicitly associate the task with a set of conditions that we take to be normal for its performance and assessment'. To model knowledge, the task is to form a true belief (about a certain question) using a given method. Then Safety says that an agent knows p via method m at w only if p is true at all sufficiently normal worlds where the agent believes p via m.

Here's an example that normality models explain. There's a trillion-sided die. You can't know whether it will land 1017 before looking. But if it comes up 1017, you can know it did by looking. Yet the risk of hallucinating is greater than one trillion. Before looking, any roll of the die is equally normal; but after looking, 1017 is more normal than other outcomes, because hallucination is abnormal.[12]

There are various conceptions of normality in the literature. On one conception, an agent knows p only if her state carries the information that p. A state carries the information that p only if normally p is true when an agent is in that state (in other words, p is true in all normal worlds where the agent is in that state).[13] Here, the normality of a state has something to do with whether it is functioning optimally. Other conceptions of normality, such as that in Carter [2019], appeal more directly to ordinary judgments about the normality of various situations. At any rate, our arguments appeal only to structural properties of various normality models, staying neutral on how exactly that notion is interpreted.

The easy possibilities are those that are sufficiently normal. But what is sufficient normality? Beddor and Pavese [2019] say the easy possibilities are those that are at least as normal as the actual world. In a good case where the actual world is very normal, the agent knows a lot, because few possibilities are as normal. In a bad case where the actual world is very abnormal, the agent knows little, because many possibilities are as normal.

---

[11]See for example [Dretske 1981; Greco 2014; Stalnaker 2015; Goodman and Salow 2018; Carter 2019; Beddor and Pavese 2019; Carter and Goldstein Forthcoming].
[12]See [Nelkin 2000; Smith 2010; Carter 2019] for a similar example involving a computer screen.
[13]See for example [Stampe 1977; Dretske 1981; Millikan 1984;Stalnaker 1999].

A second model more in the spirit of Williamson [2013] appeals to significant differences in normality. One world is significantly less normal than another when its degree of normality is quite a bit lower than the other. Then the easy possibilities at a world are those that are not significantly less normal than that world.

Goodman and Salow [2018] synthesize these two models. Assume there is always at least one most normal world. Then the easy possibilities are those that are neither significantly less normal than the most normal world(s) nor any less normal than the actual world.[14]

All three analyses are offered as models of knowledge, not just easy possibility.[15] An agent knows p just in case p is true in in any sufficiently normal world where the agent believes p.[16] These models differ in their structural properties. The first and third models predict knowledge is luminous, so that whenever an agent knows p, she also knows that she knows p (see [Williamson 2000] for criticism). When an agent knows p at w, p is true throughout the worlds v that are at least as normal as w. At any of those worlds v, p is still true throughout the worlds u that are at least as normal as v. So at any such world v, the agent knows p. So at w, the agent knows that she knows p. By contrast, the second model invalidates luminosity. Suppose p is true throughout the worlds a little less normal than w, but false at much less normal worlds. Then the agent can know p at w without knowing p at v, a world a little less normal than w. So she can know p without knowing that she knows p.

Our main observation here is that the analogue of Counterfactual Closure for normality is a non-starter:

Normic Closure If there is a sufficiently normal world where p, and if p were the case, then q would be the case, then there is a sufficiently normal world where q.

No matter how we understand sufficient normality, Normic Closure is false. Alex opens the trap door, and sees a tiger. Before Alex opens the door, it is equally normal for the tiger to be in the room or not. So there is an equally normal world where Alex opens the door and sees no tiger. In the world where there's no tiger, Alex spends the night in the room. So there is a no less normal world where Alex spends the night in the room. Actually, if Alex spent the night in the room, Alex would have been eaten by a tiger. But the world where Alex is eaten by a tiger is significantly less normal than actuality, and so is not sufficiently normal.

---

[14] Goodman and Salow also include a dimension of evidence sensitivity we suppress, since it does not affect our arguments. Instead of the most normal world, Goodman and Salow rely on the most normal world consistent with the agent's evidence. Similarly, for us what matters is not the most normal worlds in all of modal space, but the most normal worlds where an agent believes p by method m. Finally, a fourth proposal, discussed by Stalnaker [2015] and Carter [2019], is that the easy possibilities are those that aren't significantly less similar than any world that isn't significantly less normal

[15] These models all involve certain idealizations, notably the assumption that one always knows p when one is in a position to know p.

[16] As usual, our discussion could be adapted to the setting of method relativity.

If a normality model of knowledge is correct, then the failure of Normic Closure implies the failure of Counterfactual Closure. This gives a simple explanation for what has gone wrong. Counterfactuals with normal antecedents can have abnormal consequents, because counterfactuals are sensitive to similarity rather than normality.

5. Restricted Counterfactual Closure

Normality models of knowledge correctly predict that Counterfactual Closure is false. But these models undergenerate. They invalidate an important weakening of Counterfactual Closure. This weakened principle concedes that not every easy possibility that counterfactually implies q thereby makes q an easy possibility. But the restricted principle identifies an important class of easy possibilities that have this property.

Recall our earlier sufficient condition for easy possibility. Suppose I am about to flip a fair coin, and I know this. Paradigmatically, there is an easy possibility where the coin lands heads, and one where the coin lands tails. Exactly such easy possibilities explain why I am ignorant of the outcome. Now suppose some other claim q counterfactually depends on the outcome of the coin flip. Suppose for example that I've bet some money on the outcome, and I would win only if the coin were to land tails. In that case, my ignorance of the flip's outcome explains my ignorance of the bet's outcome.

We suggest the following restricted version of Counterfactual Closure, where the relevant agents knows that c is a fair coin that is about to be flipped:

Restricted Counterfactual Closure Suppose that if c were to land tails, then p would be the case. Then there is an easy possibility where p.

Again say a coin is 'active' if the relevant agents know it is fair and about to be flipped. RCC suggests that you don't know p when p would have been false if an active coin would land tails. If you are an active flip away from being wrong about p, then you don't know p.[17]

The argument against counterfactual miracles in [Goodman 2015] doesn't strictly require Counterfactual Closure. It can use Restricted Counterfactual Closure (in fact, [Goodman 2015] uses conditionals about coins in its presentation). Suppose we're in a deterministic, Newtonian world and Newtonian mechanics along with the initial conditions imply that the fair coin I am

---

[17]Most will agree that when a coin is about to be flipped, there is an easy possibility where it lands tails. Those who accept this principle will likely accept a version of Restricted Counterfactual Closure which says that p is easily possible whenever p would be the case if c were to land tails, regardless of whether the relevant agents know that c is about to be flipped. However, we'll see soon that Goodman and Salow 2018 deny this principle. They claim that sometimes an agent can know a fair coin won't land tails, even though it is about to be flipped. In their model, this happens only when the agent doesn't know whether the coin will be flipped. For their sake, we limit Restricted Counterfactual Closure to active coins.

about to flip will land heads. Then the defender of counterfactual miracles will accept that if the coin had landed tails, Newtonian mechanics would have failed.[18]

RCC puts pressure on normality models of knowledge. Only certain kinds of normality models can validate RCC, and only if they accept an interesting new principle of 'Counterfactual Contamination' connecting counterfactuals and normality. To see the point, we explore in greater detail the epistemic predicaments generated by chance events like a coin toss.

[Dorr, Goodman, and Hawthorne 2014] imagine a fair coin is flipped between 1 and 1000 times. The moment the coin lands heads, the coin stops being flipped. Possible outcomes include H, TH, TTH, …, until a series consisting of 1000 tails. Represent worlds by the number of flips they contain: world 1 is H, world 2 is TH, and so on. Imagine we are in world 1, where the coin is flipped once and lands heads. Plausibly, in this world the agent knows before observing any flips that the coin will not be flipped a thousand times.[19] At world 1, world 1000 is ruled out.

---

[18] One natural generalization of Restricted Counterfactual Closure is false. Consider a general principle connecting sufficient objective chance to easy possibility. Say there is a sufficient chance of p when the objective chance of p is greater than some threshold, say .5. Then:

Chance Counterfactual Closure If there is a sufficient chance of p and if p were the case, then q would be the case, then there is an easy possibility where q.

Some of our counterexamples to Counterfactual Closure also falsify Chance Counterfactual Closure. Imagine again that the relevant triggering event is controlled by a fair coin. For example, imagine the tiger's location is dependent on a coin flip. Then the chance of the tiger being in the room is .5, and so the chance that the agent stays in the room is .5. Nonetheless, the agent can know he won't be eaten by a tiger, even though he would be eaten by a tiger if were to stay in the room. These counterexamples similarly falsify a restriction of Chance Counterfactual Closure to cases where p is known to have a sufficient chance.

On the other hand, our counterexamples do not affect Restricted Counterfactual Closure. Even though the tiger's location is counterfactually dependent on the outcome of the coin flip, it is not true that the agent would have been eaten by a tiger if the coin landed tails. Rather, if the coin had landed tails, the agent wouldn't have stayed in the room.

Another problem for Chance Counterfactual Closure comes from the fact that we can know claims about the future which have a less than certain chance. Take a massive disjunction of the negation of all such claims about the future which the relevant agent could not easily have failed to believe. This disjunction plausibly has high chance and moreover counterfactually implies itself. So Chance Counterfactual Closure implies that there is an easy possibility where this disjunction is true. But if the disjunction is true at an easy possibility, then at least one of its disjuncts is too. But in that case the agent would fail to know that the disjunct is false, contrary to our supposition. To summarize, Chance Counterfactual Closure threatens skepticism about knowledge of the future.

[19] See [Dorr, Goodman, and Hawthorne 2014] for an argument that this claim's denial leads to widespread skepticism. To bolster their case, here is an argument from [Bacon 2014]. Consider an agent who initially doesn't know whether they are holding a fair coin or a double tailed coin. Plausibly they can know they are holding a double tailed coin if they observe a thousand tails in

Given the right counterfactual setup, this assumption surprisingly conflicts with Restricted Counterfactual Closure.

Imagine we are at world H, where the coin is flipped once and lands heads. But now imagine the world is very special. By fluke accident, the following counterfactual is true: if the coin had landed tails at least once, it would have landed tails a thousand times. Restricted Counterfactual Closure implies the agent does not know the coin won't land tails a thousand times. At the initial time, the agent knows the coin is about to be flipped. If it had landed tails at least once, it would have landed tails a thousand times. So there is an easy possibility where the coin lands tails a thousand times. So the agent doesn't know after all that the coin won't land tails a thousand times.[20]

An opponent might deny that the counterfactual if the coin landed tails at least once, then it would have landed tails a thousand times is ever true, at least in a world with a fair coin which landed heads the first time. The opponent might appeal to ties in similarity. They might say that each world 2-1000 is equally similar, and so if the coin landed tails at least once, it would have been flipped exactly twice, or exactly three times, etc...but for each particular number it's false it would have flipped exactly that many times.

This response relies on a failure of the rule of Conditional Excluded Middle, that either if p then q or if p then not q. So this response is unavailable to Stalnaker [1968] and other defenders of CEM (including [Hawthorne and Dorr forthcoming]).[21] This response is also insufficiently general. Below we identify a general problem concerning counterfactuals of the form: if the coin had landed tails, then p would have occurred, where p is abnormal. It's not clear how ties in similarity could falsify every counterfactual of this form.

Our interest is how RCC affects the normality theory. Which conceptions of sufficient normality allow Restricted Counterfactual Closure? To apply the notion of sufficient normality, we need an

---

a row. But if the agent doesn't know that a fair coin will not come up tails a thousand times, how could the agent use this data to eliminate the fair coin hypothesis?

[20] At this world, there need not be a significantly high chance that the coin will land tails a thousand times. More generally, the following principle is false: if there is a significant chance of p, and if p were the case then q would be the case, then there is a significant chance of q. This principle is refuted by Alex and the tiger. Prior to the coin flip, there is a significant chance that Alex will stay in the room. But there is not a significant chance that he will be eaten by a tiger, which is counterfactually implied by his staying in the room.

[21] A defender of CEM might respond that the uniquely most similar world where a coin is flipped at least twice is a world where it is only flipped twice. But this view implausibly requires that at every world where the coin lands heads, it is true that if a coin had been flipped twice, it would have come out TH. This is quite a radical solution to the problem. It seems to imply that before the game begins, the defender of CEM should be 75% confident that if the coin were flipped twice, it would come out TH. He is 50% confident the first flip will come out heads, and 50% confident it will come out tails. In the first case, the defender claims that it is guaranteed that if the coin were flipped twice, it would come out TH. In the second case, strong centering implies that 50% of the worlds are those where if the coin were flipped twice, it would come out TH; and 50% of the worlds are those where if the coin were flipped twice, it would come out TT.

ordering of worlds with respect to normality, and a notion of significantly less normal. Goodman and Salow 2018 model normality in this example via objective chance. Each world is denoted by its number of coin flips. World m is more normal than world n ($n < m$) just in case the chance of m is greater than the chance of n. For example, world 2 (TH) is more normal than world 4 (TTTH) because the chance of (TH) is $.5^2$, while the chance of TTTH is $.5^4$. World m is significantly more normal than world n ($n << m$) just in case in world m, the chance of having at least n coin flips is never greater than some threshold (we let this threshold be .25). For example, world 2 (TH) is significantly more normal than world 5 (TTTTH) because at world 2 the chance of the coin flipping at least 5 times is never more than $.5^4$. (After the first flip in world 2, this is the chance of flipping at least 5 times.) Generalizing, any world m is significantly more normal than another n just in case the number of flips $m + 2$ is less than n.

In the ordering above, normality supervenes on the chances of coin outcomes, as determined by the number of coin flips in the outcome. We have enriched worlds with further counterfactual structure, imagining a case where in world H the following counterfactual is true: if the coin had landed tails at least once, it would have landed tails a thousand times. This is a counterfactual abnormality: if the coin had landed tails, the world would have been abnormal. We now explore how counterfactual abnormality relates to actual abnormality.

We begin by assuming that counterfactual abnormality has no bearing on actual abnormality. For the moment, we assume that even in worlds with such counterfactuals, normality is controlled merely by the chances. We'll see that this leads to failures of RCC on all three models of normality. So suppose again we are at H, but that if the coin had landed tails at least once, then it would have landed tails a thousand times. Restricted Counterfactual Closure implies that world 1000 is sufficiently normal. Holding fixed the normality ordering from [Goodman and Salow 2018], all three models reject this conclusion.

First, consider the model from Beddor and Pavese [2019], where a world is sufficiently normal when it is as normal as actuality. In that case, at a world with exactly m flips, the easy possibilities are those where the coin flips at most m times. At TTTH the easy possibilities are H, TH, TTH, and TTTH. At H, the easy possibilities are merely H. So world 1000 is not sufficiently normal at world H, regardless of the counterfactual abnormalities H may contain.[22]

Now consider the model inspired by Williamson [2013], where a world is sufficiently normal when it is not significantly less normal than actuality. Now at a world with exactly m flips, the easy possibilities are those where it flips at most $m + 2$ times. So at TTTH the easy possibilities are H, TH, TTH, TTTH, TTTTH, and TTTTTH. When the coin flips 4 times, the agent knows that it will flip at most 6 times. At H, the easy possibilities are H, TH, and TTH. Again, world 1000 is abnormal at H.

[22][Beddor and Pavese 2019] do not apply their theory to the coin paradigm, and so might well reject the underlying normality ordering in play. Our goal here is merely to illustrate the structural features of their account.

Finally, consider the model in Goodman and Salow [2018], where a world is sufficiently normal when it is not significantly less normal than the most normal world and it is at least as normal as actuality. Then at a world with exactly m flips, with m less than 3, the easy possibilities are those where it flips at most 2 times. At any other world, the easy possibilities are those where it flips at most m times. So at TTTH, the easy possibilities are H, TH, TTH, and TTTH. At H the easy possibilities are H, TH, and TTH. Again, at H world 1000 is excluded.[23]

6. Counterfactual Contamination

Similarity models of knowledge validate Counterfactual Closure, and so they are too strong. Extant normality models of knowledge invalidate Restricted Counterfactual Closure, and so they are too weak. To resolve this impasse, we now offer a new conception of normality.

To validate Restricted Counterfactual Closure, we can modify the normality ordering so that it is sensitive to counterfactual abnormality. On this proposal, the abnormality of a world depends on how abnormal things would be given various counterfactual assumptions. Counterfactual abnormality makes for some degree of actual abnormality.

We consider two principles of ‘Counterfactual Contamination’. Each says that counterfactual abnormality at a world contaminates that world with some degree of actual abnormality. Again let c be a fair coin that the relevant agent knows is about to be flipped.[24]

Strong Counterfactual Contamination Suppose that if c were to land tails, then things would be abnormal to at least degree n. Then things are already abnormal to degree n.

Weak Counterfactual Contamination Suppose that if c were to land tails, then things would be abnormal to at least degree n. Then things are already not significantly more normal than n.

The Contamination principles say that if you’re an active flip away from significant abnormality, then you are already abnormal. Consider a world w where if c were to land tails, then worlds v or u would obtain. Suppose v is more normal than u. Weak Counterfactual Contamination implies that w is not significantly more normal than v. Strong Counterfactual Contamination implies that w is not any more normal than v.

---

[23]We pause to observe one strange feature of this model. In TTTH, Goodman and Salow [2018] predict that if the agent receives no evidence about the coin flips, then at the moment before the final flip the agent knows the coin will not land tails, although they don’t know that the coin will be flipped. The key is that any world with more coin tosses than TTTH is both significantly less normal than H and also less normal than TTTH. So the agent in TTTH knows that there are at most 4 tosses. For this reason, Goodman and Salow admit cases where a coin is about to flipped and yet there is no easy possibility that it lands tails.Such cases exist (at higher numbers of flips) even if we suppose that the coin is biased .99 in favor of tails.

[24]Of course, generalizing these contamination principles to domains beyond coins is a delicate matter. Our remarks in the conclusion may offer some initial guidance here.

To illustrate Contamination, consider how we judge the normality of people. Imagine Jeremy lives a perfectly normal life, never losing his temper very much. Imagine one day Jeremy bets 'heads' on a coin he knows is about to be flipped. Actually, the coin will land heads and Jeremy will be slightly pleased. But counterfactually, if the coin had landed tails then Jeremy would have had a complete meltdown. He would have behaved completely erratically. Weak Counterfactual Contamination says that even though Jeremy doesn't behave abnormally, he is still somewhat abnormal for satisfying this counterfactual. Strong Counterfactual Contamination says that actual Jeremy is just as strange as he would have been had he had his meltdown.

We can imagine different versions of this case, depending on the basis for the abnormal counterfactual. In the most poignant version, we can imagine the categorical basis for the counterfactual is a statistical mechanical accident. Some slightly freak micro-realization of a fairly mundane macrostate generates a counterfactual propensity to strange behavior on this one occasion; but not in a way that is correlated with Jeremy's usual response to stimuli. For most bets, Jeremy wouldn't have had a meltdown had he lost. Only this one bet, we can imagine, would trigger a counterfactual meltdown.

Now consider how Contamination improves each model of normality. Start with the model from Beddor and Pavese, where the easy possibilities are those at least as normal as actuality. When supplemented with Strong Counterfactual Contamination, this model validates Restricted Counterfactual Closure. The normality of a world with exactly one coin flip depends on what counterfactuals are true there. A world with exactly one coin flip where extra coin flips would have counterfactually implied a thousand coin flips is just as abnormal as a world that already has the thousand coin flips. So an agent in the former world cannot rule out being in the latter world. Interestingly, this model does not validate RCC when supplemented with Weak Counterfactual Contamination. Now a world with exactly one coin flip where any more coin flips would counterfactually imply a thousand coin flips can be less abnormal than a world that already has the thousand coin flips, and so the latter will not be easily accessible from the former in Beddor and Pavese's model.

Now consider the model of normality inspired by Williamson [2013], where a world is sufficiently normal when it is not significantly less normal than actuality. Here, Weak Contamination is sufficient to ensure Restricted Counterfactual Closure. Consider again the world with exactly one flip, where more flips counterfactually imply a thousand. Weak Contamination implies that the world with a thousand flips is not significantly less normal than the base world. So there is an easy possibility with a thousand flips.

Finally, consider the model from Goodman and Salow, where the easy possibilities are those not significantly less normal than the most normal world, and those at least as normal as the actual world. Strong Contamination validates Restricted Counterfactual Closure within this framework, for the same reason as in Beddor and Pavese.

As with Beddor and Pavese's model, this theory does not validate RCC given only Weak Contamination. Consider three worlds. At Ordinary Heads, the coin lands heads and no strange counterfactuals are true. At Strange Heads, the coin lands heads, but if the coin had landed tails, it would have landed tails a thousand times. At One Thousand, the coin lands tails a thousand times. We can show there is a violation of RCC at Strange Heads. First, we've assumed that One Thousand is significantly less normal than Ordinary Heads, which is among the most normal worlds. Second, if we reject Strong Contamination, it could be that One Thousand is less normal than Strange Heads. So suppose this is the case. Now One Thousand is significantly less normal than the most normal worlds, and less normal than Strange Heads. It follows that One Thousand is epistemically inaccessible from Strange Heads. This contradicts RCC. At Strange Heads, the agent knows that the coin is about to be flipped. If the coin were to land tails, then One Thousand would be actual. Yet the agent can rule out One Thousand, and so One Thousand is not an easy possibility.

Which Contamination principle should we accept, weak or strong? We think Strong Contamination is too strong. Counterfactual abnormality is relevant to actual normality; but it isn't that relevant. Recall Jeremy's counterfactual meltdowns. Strong Contamination implies that Jeremy, who never has a real meltdown and only would have by statistical mechanical fluke, is weirder than someone who actually has a slightly less intense meltdown. Similarly, consider coins. In the one world, the coin is only flipped once, but would have flipped a thousand times if it had flipped even twice. In the second world, the coin actually flips exactly 999 times. Strong Contamination implies the first world is less normal than the second.

Here is an even more forceful argument against Strong Contamination.[25] Imagine a case where 1000 independent fair coins are about to be tossed simultaneously, and Danny knows this. Label the coins 1 through 1000. As a matter of fact, Danny is in world 471, where coins 1-529 will land heads and coins 530-1000 will land tails (so that 471 total coins land tails). For similar reasons to the above, Danny can know the coins will not all land tails. But considerations of causal independence familiar from Morgenbesser's coin suggest that if any given coin had landed differently, the other coins would have landed the same.[26] So in world 471, the following counterfactual is true: if coin 529 had landed tails, then coins 529-1000 would have landed tails. Strong Contamination implies that world 471 is equally normal as world 472, where coins 529-1000 land tails. Similarly, in world 472, the following counterfactual is true: if coin 528 had landed tails, then coins 528-1000 would have landed tails. Strong Contamination implies that world 472 is equally normal as world 473, where coins 528-1000 land tails. Repeating this process, we reach the result that world 471 is equally normal as world 1000, where coins 1-1000 all land tails. For this reason, Strong Contamination implies that the agent in world 471 doesn't know whether all 1000 coins will land tails. In this way, Strong Contamination leads to skepticism.

---

[25]Thanks to Cian Dorr.

[26]For example, imagine Danny places separate bets for 'heads' on each coin. It is plausible that if coin 529 had landed tails, Danny would have lost his bet on coin 529, but would still have won his bet on coins 1-528, and would have still lost his bet on coins 530-1000. For discussion see among others [Slote 1978; Schaffer 2004].

There is reason to care which contamination principle is correct. Theories of normality which require Strong Counterfactual Contamination characteristically lead to the luminosity of knowledge. By contrast, theories which make do with only Weak Counterfactual Contamination lead to anti-luminosity. We can generalize this point beyond the normality models we've considered so far. We now show that Safety and RCC lead to a failure of luminosity in any theory which accepts a generalization of Weak Contamination without Strong Contamination.

If Weak Contamination is true and Strong Contamination is false, then worlds with counterfactual abnormality are less abnormal than worlds with actual abnormality. Think of the degree of normality as determining the distance from a given world. Worlds with actual abnormality are further away from worlds that simply would have had that abnormality had an active coin landed tails. If we depart from normality models, we can replace this notion of abnormality with the more general notion of epistemic accessibility. If world w is closer than world v in epistemic distance, then there could be an agent who could rule out w without being able to rule out v.

To express this idea more precisely, we move from the coin game to a second paradigm where we have crisp judgments about relative ease of knowability. Return to our case of Jeremy and his counterfactual meltdown. Order meltdowns in strength from 0 to 1000, with 1000 the highest level of meltdown. As before, we imagine a series of cases where Jeremy bets on an active coin and if the coin had landed tails, he would have lost and had an n meltdown, for some strength n. If we accept Weak Contamination without Strong Contamination, we will think that it is easier to rule out Jeremy having an n meltdown than it is to rule out a case where Jeremy would have had an n meltdown if the coin had landed tails. That is:

Counterfactual Ignorance For any n, if the strongest proposition known is that Jeremy won't have an n meltdown, then for all that is known, if the active coin were to land tails, Jeremy would have had an n meltdown.

We now show that the combination of RCC and Counterfactual Ignorance leads to failures of luminosity.

First, we suppose that there is a strongest proposition known about Jeremy's meltdowns. For example, imagine that the agent knows that Jeremy won't have an 800 meltdown; but for all the agent knows, Jeremy will have a 799 meltdown. Now consider three worlds. In 0-0, Jeremy has a degree 0 meltdown (this is no meltdown at all), and Jeremy would have had a degree 0 meltdown if the coin had landed tails. In 0-800, Jeremy has a degree 0 meltdown, but Jeremy would have had a degree 800 meltdown if the coin had landed tails. Finally, in 800-800, the coin lands tails and Jeremy does have the degree 800 meltdown.

By construction, the agent in 0-0 knows they aren't in 800-800, since they can rule out Jeremy actually having the 800 meltdown. By Counterfactual Ignorance, the agent in 0-0 doesn't know whether they are in 0-800. Since 800 is the lowest degree actual meltdown they can rule out, they can't rule out a counterfactual meltdown of 800. Finally, RCC implies that in 0-800 the agent can't rule out being in 800-800. After all, in 0-800 the agent knows the coin is about to be flipped. And in 0-800, if the coin were to land tails, then 800-800 would be actual. So RCC implies that 800-800 is an easy possibility at 0-800, and so Safety implies that 800-800 is epistemically possible at 0-800. (Throughout this argument we suppose that the agent's beliefs are insensitive to the meltdown, so that the agent at 800-800 would still have believed that Jeremy will not have an 800 meltdown. After all, we suppose Jeremy's meltdown will come about by statistical mechanical fluke, unforeseeable by daily observation of his character.)

In summary, we've shown that Safety and Restricted Counterfactual Closure lead to failures of luminosity when paired with a conception of counterfactual knowledge implicit in any normality model that accepts Weak Contamination without Strong Contamination. In this way, our points generalize beyond the setting of normality models.[27]

7. Counterfactual Triggers

We've shown Restricted Counterfactual Closure has significant consequences for models of normality. But our discussion so far has a noticeable gap. We've focused on counterfactual antecedents about the outcomes of fair coin flips. But what is it about coin flips that make them so relevant to knowledge? One natural thought appeals to normality. Either outcome of a coin flip is normal. Whenever p is normal and p counterfactually implies q, there is an easy possibility where q. Yet this natural thought is untenable. It is tantamount to the Normic Closure principle rejected above. Recall that it is normal for the tiger to be in the room, and the tiger's presence counterfactually implies Alex's death; and yet there is no easy possibility in which Alex dies.

Call a 'counterfactual trigger' any proposition p where q is an easy possibility whenever it is counterfactually implied by p and p is an easy possibility. We've argued that counterfactual triggers lead to counterfactual contamination. Non-actual abnormal events that counterfactually depend on non-actual counterfactual triggers make for actual abnormality. So far, we've claimed

[27] A referee notices that certain ideas about chance produce potential complications. Suppose for example that the chance of a coin toss depends on the frequencies of heads and tails among coins in the world. A certain coin is fair and comes up tails, but if it had come up heads, the frequencies would have determined that the chance of heads was very low. The referee suggested that the world in which the coin comes up heads is way more abnormal than the world where the coin comes up tails fairly. If knowledge is understood in terms of normality, one could then know that the heads world will not obtain. One response here is to reject the possibility of undermining chances. Many have tried to fix the best system account of chance (as in [Lewis 1980]) to avoid undermining chances. But it is also worth noting that even granting this surprising picture of chance, it is not clear that there is a counterexample to RCC here. The principle of Counterfactual Contamination rules out just such a counterexample. If we were really a coin flip away from being in such an abnormal scenario, then by Counterfactual Contamination the actual world is already abnormal enough to make the bizarre abnormal an easy possibility.

that coin flips are counterfactual triggers. In this conclusion, we suggest that counterfactual triggers can be understood in terms of macroscopic regularity.

The world is rife with macroscopic causal regularities. Tigers eat people. People avoid being eaten. People don't like to eat nails. People die when they eat nails. Normality is related to macroscopic regularity.

Counterfactual possibilities often but not always violate macroscopic regularities. Again suppose there is a tiger in the room. If Alex had entered the room, he would have been eaten by the tiger. When we counterfactually suppose that Alex entered the room, we consider a world in which there is a tiger in the room, and Alex enters the room anyways. This violates the macroscopic regularity that people who fear danger avoid danger. But the counterfactual imagination tolerates this violation.

Many have observed that counterfactuals are often evaluated relative to causal structure. In the tiger example, we have a chain of events: the tiger being trapped in the room; Alex refraining from entering the room; Alex surviving the day. Each event in the chain causes the next one. Counterfactuals characteristically intervene in causal chains.[28] The above counterfactual holds fixed that the tiger is trapped in the room, and severs the macroscopic regularity linking the tiger's presence to Alex's absence. This introduces abnormality into the world.

Coin flips are special because they don't require causal intervention on the macroscopic laws. Granted, in a deterministic world, my coin's landing heads is causally necessitated by the initial conditions of the universe. But at the macroscopic level, there are no facts about the time before the coin flip that imply the coin will land heads. Coin flips are microscopically deterministic, but macroscopically indeterministic.

We imagined that if the coin had been flipped twice, it would have been flipped a thousand times. A thousand flips is quite an abnormal outcome. But it is a different kind of abnormality. In this world, no causal intervention on the macroscopic laws is required. No macroscopic regularity need be violated.

We suggest that counterfactual triggers are easy possibilities that could obtain without macroscopic intervention. The absence of macroscopic intervention is a counterfactual trigger because macroscopic causal interventions are dramatically abnormal. There are many kinds of abnormality, but this one is especially repugnant.[29]

---

[28]For discussion of how to integrate this idea into the theory of counterfactuals, see among others: [Hiddleston 2005; Briggs 2012; Santorio 2019].
[29] Thanks to Sam Carter, Cian Dorr, Jeremy Goodman, Zach Goodsell, Dan Waxman, and the audience at the Lingnan University Department of Philosophy Seminar Series for feedback on this paper.

REFERENCES

Bacon, Andrew 2014. Giving Your Knowledge Half a Chance, *Philosophical Studies* 171/2: 373-397.

Beddor, Robert and Pavese, Carlotta 2019. Modal Virtue Epistemology, *Philosophy and Phenomenological Research* 101/1: 61-79.

Bennett, Jonathan 1984. Counterfactuals and Temporal Direction, *Philosophical Review* 93/1: 57-91.

Briggs, R.A. 2012. Interventionist Counterfactuals, *Philosophical Studies* 160/1: 139-166.

Carter, Sam 2019. Higher Order Ignorance Inside the Margin, *Philosophical Studies* 176/7: 1789-1806.

Carter, Sam and Goldstein, Simon Forthcoming. The Normality of Error, *Philosophical Studies*.

Dretske, Fred 1981. *Knowledge and the Flow of Information*, Cambridge, MA: The MIT Press.

Dorr, Cian 2016. *Against Counterfactual Miracles*, *Philosophical Review* 125/2: 241-286.

Dorr, Cian, and Goodman, Jeremy, and Hawthorne, John 2014. Knowing Against the Odds, *Philosophical Studies* 170/2: 277-87.

Dorr, Cian, and Hawthorne, John Forthcoming. *Conditionals*, Oxford: Oxford University Press.

Goldstein, Simon and Waxman, Dan 2020. Losing Confidence in Luminosity, *Noûs* doi: 10.1111/nous.12348.

Goodman, Jeremy 2015. Knowledge, Counterfactuals, and Determinism, *Philosophical Studies* 172/9: 2275-8.

Goodman, Jeremy and Salow, Bernhard 2018. Taking a Chance on KK, *Philosophical Studies* 175/1: 183-96.

Greco, Dan 2014. Could KK be Ok?, *The Journal of Philosophy* 111/4: 169-197.

Hiddleston, Eric 2005. A Causal Theory of Counterfactuals, *Noûs* 39/4: 632-657.

Lewis, David 1973 *Counterfactuals*, Cambridge, MA: Harvard University Press.

Lewis, David 1979. Counterfactual Dependence and Time's Arrow, *Noûs* 13/4: 455-476.

Lewis, David 1980. A Subjectivist's Guide to Objective Chance, in *Studies in Inductive Logic and Probability*, ed. Richard Jeffrey, Berkeley: University of California Press: 263–293.

Manley, David 2007. Safety, Content, A Priority, Self-Knowledge, *The Journal of Philosophy* 104/8: 403-423.

Millikan, Ruth 1984. *Language, Thought, and Other Biological Categories*, Cambridge: MIT Press.

Nelkin, Dana K. 2000. The Lottery Paradox, Knowledge and Rationality, *Philosophical Review* 109/3: 373-409.

Santorio, Paolo 2019. Interventions in Premise Semantics, *Philosophers' Imprint* 19/1: 1-27.

Smith, Martin 2010. What Else Justification Could Be, *Noûs* 44/1: 10-31.

Sosa, Ernest 1999. How to Defeat Opposition to Moore, *Philosophical Perspectives* 13: 141-153.

Schaffer, Jonathan 2004. Counterfactuals, Causal Independence and Conceptual Circularity, *Analysis* 64/4: 299-309.

Slote, Michael 1978. Time in Counterfactuals, *Philosophical Review* 87/1: 3-27.

Stalnaker, Robert 1968. A Theory of Conditionals, *Studies in Logical Theory, American Philosophical Quarterly*, Monograph: 2, 98-112.

Stalnaker, Robert 1999. *Context and Content: Essays on Intentionality in Speech and Thought*, Oxford: Oxford University Press.

Stalnaker, Robert 2015. Luminosity and the KK Thesis, In *Externalism, Self-Knowledge, and Skepticism*, ed. Sandy Goldberg, Cambridge: Cambridge University Press.

Stampe, D 1977. Towards a Causal Theory of Linguistic Representation, *Midwest Studies in Philosophy* 2/1: 42-63.

Williamson, Timothy 2000. *Knowledge and its Limits*, Oxford: Oxford University Press.

Williamson, Timothy 2009. Probability and Danger, *The Amherst Lecture in Philosophy* 4, 1-35.

Williamson, Timothy 2013. Gettier Cases in Epistemic logic, *Inquiry* 56/1: 1-14.