ARTICLE

Noûs

'Logic will get you from A to B, imagination will take you anywhere'

Francesco Berto[1,2]

[1]University of St Andrews, St Andrews, United Kingdom

[2]University of Amsterdam, Amsterdam, The Netherlands

Correspondence
Francesco Berto, University of St Andrews, St Andrews, United Kingdom and University of Amsterdam, Amsterdam, The Netherlands.
Email: fb96@st-andrews.ac.uk

Abstract
There is some consensus on the claim that imagination as suppositional thinking can have epistemic value insofar as it's constrained by a principle of minimal alteration of how we know or believe reality to be – compatibly with the need to accommodate the supposition initiating the imaginative exercise. But in the philosophy of imagination there is no formally precise account of how exactly such minimal alteration is to work. I propose one. I focus on counterfactual imagination, arguing that this can be modeled as simulated belief revision governed by Laplacian imaging. So understood, it can be rationally justified by accuracy considerations: it minimizes expected belief inaccuracy, as measured by the Brier score.

1 | THE PUZZLE OF IMAGINATIVE USE

The spurious Einsteinian quote making for our title evokes the virtues of lateral thought. But the logician may retort: at least logic will take you from your given *A* to the proper *B*s, namely those that follow logically from *A*. Imagination, instead, is a runabout inference ticket: like *tonk*, it may take you from any *A* to any *B* whatsoever, and so to no epistemically valuable place. This links to what Amy Kind and Peter Kung call 'the Puzzle of Imaginative Use' in their introduction to the wonderful collection *Knowledge Through Imagination* (Kind and Kung 2016): given that imagination is escape from reality, arbitrary in ways belief is not (one can easily imagine, but one can't easily make oneself believe, that all of Oxford has been painted pink), how can it have epistemic value? How can it give us accurate beliefs on reality?

We use 'to imagine' to refer to lots of activities: daydreaming, hallucinating, entertaining, mental wandering via free associations of ideas. But when one focuses on propositional imagination (imagining that one jumps over a stream; that Arif is in the kitchen; that Oswald hasn't killed Kennedy), of the kind involved in suppositional thinking,[1] one can extract from the literature in the philosophy of imagination some consensus on the beginning of an answer to the Puzzle.

First, a number of authors (e.g., Currie and Ravenscroft, 2002; Goldman, 2002b) agree on a simulationist view: imagination recreates counterparts non-imaginative states. Propositional imagination in particular is often taken as simulating *belief* (Arcangeli (2019) calls it 'cognitive imagination'). Second, such imagination can have epistemic value insofar as it's *constrained* (Kind, 2016; Williamson, 2016; Langland-Hassan, 2016; Badura, 2021). One may start by supposing whatever *A* one likes but, once the initial supposition is in, how one develops it in imagination is not arbitrary: it is governed by some principle of minimal alteration of how we know or believe reality to be, compatibly with the need to accommodate the supposition that *A*.

When, for instance, one imagines that one jumps over a stream, wondering whether one would make it to the other side if one tried (Williamson, 2016), it would spoil the exercise to imagine that one magically grows wings and flies over the stream. One sticks to reality as much as possible in the imagined scenario, e.g., by holding fixed one's presumed physical capacities. This is called 'reality-orientation', or 'reality-monitoring', in cognitive psychology (Johnson and Raye, 1981): we 'use our knowledge of how the world works to keep the imagining realistic' (Davies, 2019, 187). When *Wuthering Heights* tells us that Heathcliff meets Catherine for the last time, we imagine that Heathcliff is dressed as a late eighteenth century gentleman, not as a punk from the seventies, although the text says nothing on this. That's because we have certain beliefs on how people dressed at the time when the events of the story are located and, lacking explicit information to the contrary from the novel, we import our beliefs into the scenario. In general, we look at an imagined situation which is as similar as it gets to how we take reality to be, but for the initial supposition.

What principle of minimal alteration is at work here, and how exactly would such a principle confer to the activity its epistemic value? Little of formal precision can be found in the philosophical literature on imagination. There are, to be sure, psychological studies on how people tend to change their beliefs under supposition (Byrne, 2005; Kosslyn and Moulton, 2009; Davies, 2019). Insofar as these are descriptive of what people generally do, it's not straightforward to extract from them an answer to our question, which has to do, rather, with rational normativity: the answer may be compatible with people generally not living up to the standards of epistemically optimal suppositional thought, whatever these turn out to be.

The aim of this paper is to provide an answer. I exploit an idea taken from formal epistemology, which, to the best of my knowledge, has never been connected to the issue of the epistemic value of imagination. This may depend on its relying on a belief revision procedure, which has received little attention in comparison to Bayesians' favourite, conditionalization.

I first focus (section 2) on a specific kind of propositional imagination, namely counterfactual imagination. In section 3, I unpack how this differs from its counterpart, imagination 'in the indicative mood'. In section 4, I introduce the revision procedure at issue, due to Lewis and called *imaging*. I explain why imaging should be of interest to philosophers of imagination: it captures in a precise way the idea of minimal alteration, or of 'looking at the most similar

---

[1] I use 'to suppose' and 'to imagine' more or less interchangeably. An amount of literature takes the two verbs as labeling distinct mental activities (see Arcangeli (2019) for a rich discussion), e.g., on the grounds that only imagination would essentially involve mental imagery (Kind, 2001). Some authors disagree (Van Leeuwen, 2013; Williamson, 2016; Gregory, 2016; Berto, 2022). I don't think much in our discussion hangs on this issue.

scenario where *A'*, in an exercise of counterfactual imagination triggered by the supposition of *A*.

In section 5 I present the main idea: a specific kind of imaging, called *Laplacian*, is of epistemic value insofar as it minimizes the expected inaccuracy of one's degrees of belief, once this is assessed by the measure of inaccuracy that goes under the label of 'Brier score'. So I explain how the Brier score works. For the argument to be persuasive one needs to buy some assumptions on the goodness of that measure and on the epistemic value of accuracy. So I rehearse considerations from the literature on this, taken from 'accuracy-first' epistemology. This has focused mostly on the representational mental state of belief, or on credences as degrees of belief, but, I argue there, such considerations can be straightforwardly extended to counterfactual imagination.

In the coda of section 6, I wonder whether the rationality of imaging is compatible with that of Bayesian conditionalization, given how the two differ. I argue that it is, exploiting a parallel with two non-competing kinds of non-probabilistic belief revision from the AI literature. Here we go!

2 | INDICATIVE AND COUNTERFACTUAL IMAGINATION

What would be the case if something was the case? We ask questions of this kind when we look for causal explanations ('Would we see such a particle track if the atom was ionized?'), ascertain responsibilities ('Would he have hit the brakes in time, had he not been distracted?'), learn from past mistakes ('Would they have won the match, had they played with a different module?'), decide what to do next ('Would I make it to the other side if I jumped over the stream?'). We often answer by imagining a situation where some *A*, which is our suppositional input, obtains (the atom is ionized, the driver is not distracted, etc.); and we wonder what would be likely under the supposition. This is often called subjunctive or counterfactual imagination (Byrne, 2005, ch. 1). The importance of the activity can hardly be overestimated. According to eminent computer scientists like Judea Pearl (2013), AI research will only move beyond stochastic machine learning when we come up with systems capable of tracking causal connections over and above probabilistic correlations. Such systems will have to be able to imagine things being otherwise than they are, to address the question of what would happen if we brought it about that *A*, as opposed to just observing that *A*.

The use of 'if' in the phrasing of hypothetical questions links to conditionals: we often assess them by imagining a situation where the antecedent is true and estimating the chances of the consequent in the imagined scenario (Evans and Over, 2004; Williamson, 2020). One can then distinguish imagination in the indicative and counterfactual mood by mapping them to the familiar distinction between kinds of conditionals. Back to the old chestnut, which Lewis (1973) attributes to Adams:

1. If Oswald did not kill Kennedy, someone else did.
2. If Oswald had not killed Kennedy, someone else would have.

We can assess both by imagining a situation where the antecedent is true and wondering about the status of the consequent there. In both cases, we imagine that Oswald has not killed Kennedy. But we deem (1) true, (2) false. So the two imaginings must differ. How?

One mainstream answer has it that we can only imagine in the indicative mood that *A* when *A* has some chance of being true for us; whereas that's not the case for counterfactual imagination. So in his book on conditionals Bennett claims of indicatives:

> [Y]ou cannot say what the upshot is of adding to your belief system something you actually regard as having *no* chance of being true [...]. There is abundant intuitive evidence that *nobody has any use for A → B when for him P(A) = 0*. (Bennett, 2003, 55, notation adjusted)

Prominent accounts of indicatives connect their subjective probability, and so their (degree of) believability or acceptability, to the corresponding conditional probability. This links to (what we now call) the Ramsey test (Ramsey, 1990): we assess 'If A, then B' by hypothetically adding A to our belief system, adjusting it for consistency in the light of the addition, and checking the resulting status of B. So the probability of an indicative is connected to the probability of the consequent conditional on the antecedent, $P(B|A)$ (Jackson, 1979). It is generally agreed that the connection cannot be identity, after the famous triviality results provided by Lewis (1976), Hajek (1989), and others. However, psychologists of reasoning can content themselves with an empirically robust correlation: the idea is at the core of the suppositional account of indicatives due to authors like Evans and Over (2004). Now when conditional probabilities are taken as ratios of unconditional ones, $P(B|A) := P(A \wedge B)/P(A)$, they end up undefined when $P(A) = 0$, giving mathematical substance to Bennett's claim.

I beg to differ: one is quite sure that one exists but one can imagine that one doesn't and reasonably assess the indicative 'If I do not exist now, then thinking occurs without a thinker' (Williamson, 2020, came up with the example). Less philosophical cases happen all the time in mathematical reasoning. In the business of proving the infinity of primes, our teacher asserts: 'If there is a largest prime, then some number is both prime and composite'. They regard the antecedent as impossible. So do we: it's clear to us that the teacher is carrying out a proof by *reductio*.

One who replies that these are covert counterfactuals should be asked, How so? On the face of it, they are indicatives and it may seem *ad hoc* to postulate a distinction between grammatical appearance and reality just to save the view. According to Joyce:

> [I]t is often assumed that any form of probabilistic belief revision that involves 'raising the dead' by increasing the probabilities of certainly false propositions must involve counterfactual beliefs. This is not so. It is logically consistent both to be certain that some proposition is false and yet to speculate about what the world is like if one is in fact wrong. To be subjectively certain of something is, after all, not the same as regarding oneself infallible on the matter. (Joyce, 1999, 203)

Unpretentious thinkers can imagine in the indicative mood that A and consider what is likely to be the case then, also when they assign A zero chances: one can non-trivially assess the indicative (1) above even when one is certain that Oswald did kill Kennedy (Gillies, 2004).

Next, one can avoid $P(A) = 0$'s making the corresponding conditional probability undefined (or assigned an arbitrary default value): e.g., instead of reducing conditional probabilities to unconditional ones, one takes them as primitive and uses Popper functions (Van Fraassen, 1995; Arlo-Costa and Parikh, 2005; Baltag and Smets, 2008). Taking conditional probabilities as more basic than unconditional ones may even be psychologically plausible in connection to suppositional thought: we often feel comfortable in assessing the conditional probability of a consequent supposing the antecedent, without feeling positioned to assess the probability of the antecedent or that of the consequent unconditionally (Williamson, 2020, 34).

3 | COTENABLES

But if we shouldn't draw the line between indicative and counterfactual imagination the mainstream way, how *should* we draw it? Having started an exercise of hypothetical thought by imagining that *A*, we don't just stick with the explicitly given input *A*. We integrate it by importing 'cotenable' (Lewis, 1973, following Goodman) background assumptions we take as holding in the hypothetical scenario, drawn from our belief or knowledge base: 'children and adults elaborate the pretend scenarios', filling in the explicit input with 'an increasingly detailed description of what the world would be like if the initiating representation were true' (Nichols and Stich, 2003, 26-28).

Compare a case of actual belief change: it's around noon and one sees that Arif has moved to the kitchen. One then minimally revises one's beliefs compatibly with the news. In the updated belief system one will have dropped the previously held belief that Arif is in the living room; one may now find the thought that Arif is cooking likely enough to come to fully believe it's true (or anyway increase one's confidence in its truth, etc.) But one will retain a number of previously held beliefs (or degrees thereof): one will still believe that there's an oven in the kitchen, that Arif can use it to cook, etc. In suppositional thought, one does not get the news online, perceptually, e.g., by seeing that Arif is in the kitchen. One just imagines, in offline mode (Williamson, 2016), that Arif is there, thus simulates getting the news; and one checks what's likely in the imagined scenario.

Now the difference between indicative and counterfactual imagination, I submit, lies in *which* beliefs are cotenable in the two modes. When we imagine indicatively that Oswald hasn't killed Kennedy we retain our belief that Kennedy was actually killed – and so it must have been someone else, and so we find the indicative (1) acceptable. When we imagine the same thing counterfactually, we depart from how we know or believe actuality to be like, in such a way that we can make relevant comparisons with it from within the supposition: we relinquish that belief and find it plausible that nobody else kills Kennedy in the counterfactual scenario – and so we don't find (2) acceptable. Our sticking to actuality in the former case, not in the latter, is mirrored in the famously different behaviour of the two kinds of conditionals when actuality operators are embedded in their consequent: 'If Midori is working, she is actually working' is trivially true; 'If Midori was working, she would actually be working' is not.

But how exactly do we minimally alter the probabilities we assign to various claims in counterfactual imagination? We may take indicative imagination, in Bayesian spirit, as a sort of simulated belief revision governed by conditionalization. But indicative and counterfactual imagination differ in cotenability. So what is the latter governed by?

In the very same 1976 paper, in which Lewis came up with the aforementioned triviality results for the probabilities of indicatives, he also provided the tools to reply to that question: he introduced *imaging* as a form of belief change which is minimal in a different sense from conditionalization. Let's see how this works, and why it gives a natural way of making precise the idea of counterfactual imagination as simulated belief revision.

4 | IMAGING IN IMAGINATION

We need a little bit of possible worlds machinery. Say we have a finite set of possible worlds W, on which a total closeness ordering is defined in the style of the Stalnaker (1968) and Lewis (1973)

semantics for conditional logic. Closeness is naturally understood as representing comparative similarity in the relevant respects between possible worlds.

Next, different kinds of imaging come from using different *transfer functions*, specifying how probabilities should be moved around between worlds under a counterfactual supposition. A transfer function's general form is: '$T_A(w, w_1)$'. T takes as input the supposition A and two $w, w_1 \in W$ and outputs the proportion of the probability, assigned to w by a probability distribution P, which is to be moved to w_1 when we counterfactually imagine that A. We define the probability of world w_1 under the counterfactual supposition of A according to transfer function T:

$$P(w_1 |_T A) := \sum_{w \in W} P(w) T_A(w, w_1)$$

This just means: the probability assigned to w_1 under the image of A is given by taking, for each world w, the proportion of the probability of w that has to be transferred to w_1 according to the transfer function, and adding them all together. Next, we define the probability of B under the image of A just as the sum of the probabilities of the worlds where B is true under that image. Let '$|B|$' stand for the truth set of B (the set of worlds where B is true, or the possible-worlds proposition that B):

$$P(B |_T A) := \sum_{w_1 \in |B|} P(w_1 |_T A)$$

We may start by assuming, following Stalnaker's conditional logic (as Lewis (1976) did when he introduced imaging), that for each w and A there is a single closest A-world, w_A, i.e., a single world most similar to w where A is true (we'll relax the Stalnakerian assumption later on). The corresponding transfer function S goes:

$$S_A(w, w_1) = \begin{cases} 1 \text{ if } w_1 = w_A \\ 0 \text{ if } w_1 \neq w_A \end{cases}$$

The probability of each w is transferred to its closest A-world, w_A. Now w_A may be the closest world to more than one world; so one adds up the probabilities of all of those worlds. Each A-world keeps the probability it had before, and may gain probabilities transferred from non-A-worlds, i.e., worlds where A is not true. Probabilities are only moved around between worlds, but not created or destroyed. So it's easy to see that $P(|_S)$ is a probability distribution when P is.

Both imaging and conditionalization are minimal forms of revision. So they both comply with the idea of minimal alteration governing suppositional thought: when we suppose that A, we adjust our belief system just as little as possible, compatibly with accommodating the supposition of A. But they are minimal in different senses. As Lewis has it:

> Imaging P on A gives a minimal revision in this sense: unlike all other revisions of P to make A certain, it involves no gratuitous movement of probability from worlds to dissimilar worlds. Conditionalizing P on A gives a minimal revision in this different sense: unlike all other revisions of P to make A certain, it does not distort the profile or probability ratios, equalities, and inequalities among sentences that imply A. (Lewis, 1976, 311)

A simple Lewisian example shows how the two differ: say we have three equiprobable worlds w, w_1, w_2. Their probabilities must add up to 1, so the probability of each one is $1/3$; w_1 and w_2 make A true while w doesn't. The ordering has it that w_1 is closer to w than w_2. When we revise the probability assignment by conditionalizing on A, we kick out w (because A fails to be true there) and renormalize: so $P(w_1|A) = P(w_2|A) = 1/2$. Imaging, instead, exploits the similarity ordering: all of w's probability is transferred to the closest A-world w_1. Thus, $P(w_1|_S A) = 2/3$ while $P(w_2|_S A)$ stays $1/3$: w_1 gets the whole probabilistic bonus.

Now the difference between conditionalization and imaging clarifies the aforementioned difference in the beliefs we deem cotenable when we imagine in the indicative and counterfactual mood. Gärdenfors (1982) has proved that conditionalization, unlike imaging, has a property, called *conservativity*, whose natural translation in terms of imaginative thinking goes as follows: when one imagines in the indicative mood that A wondering how likely B is then, and one is certain of C for a cotenable C, i.e., $P(C) = 1$, then also $P(C|A) = 1$. For example: when one wonders what is the case if Oswald did not kill Kennedy, and one is certain that Kennedy has been killed, one will retain that certainty under the supposition and import it into the imagined scenario. This explains why our indicative (1) from the pair of Oswald conditionals is deemed acceptable.

But imaging is not conservative: it could be that C becomes uncertain for one when one imagines in the counterfactual mood that A, wondering how likely B would then be. So when one wonders what would have been the case if Oswald had not killed Kennedy, one may relinquish one's belief that Kennedy was killed and not import this into the imagined scenario. Thich explains why our counterfactual (2) from the pair of Oswald conditionals is deemed unacceptable.

Imaging should be of interest to philosophers of imagination: by re-distributing probabilities under the supposition of A while giving a bonus, so to speak, to the most similar A-worlds, it makes formally precise the aforementioned intuitive picture of what we do when we engage in an act of constrained counterfactual imagination for epistemic purposes. When we image on A wondering how likely B would then be, we focus on situations which are *closest* or *most similar* to how we take reality to be, but where A is true. This captures the idea of constrained or reality-oriented imagination described in section 1: when we counterfactually suppose that $A = We\ jump$ *over the stream*, we look at the likelihood of $B = We\ make\ it\ to\ the\ other\ side$, after boosting the probability of the worlds which are as similar as it gets to the actual one (e.g., where we retain our current physical capacities), except that we jump.

Once rephrased in precise imaging-probabilistic terms, the Puzzle of Imaginative Use for counterfactual imagination becomes: Under which conditions are we rationally justified in believing B on the counterfactual supposition of A, to the extent that we deem it likely under the image of A?

The issue is now pressing, for imaging diverges from conditionalization. There are well-known (so-called dynamic or diachronic) Dutch book (Lewis, 1999; Mahtani, 2012) as well as epistemic accuracy (Greaves and Wallace, 2006; Pettigrew, 2016) arguments, to the effect that any plan to update one's degrees of belief that diverges from conditionalization is doomed to be, respectively, Dutch-bookable (i.e., such that one can set up a set of bets against one who so plans to update, guaranteeing a secure loss), or accuracy-dominated by (i.e., guaranteed to be farther from the truth than) a plan that doesn't so diverge. Counterfactual imagination as governed by imaging seems to be in trouble.

I will not deal with Dutch book arguments for reasons of space, and also because they have already been met with growing criticism (see Titelbaum (2022), ch. 9, for a summary of the discussion). But I will deal with accuracy: under certain assumptions, accuracy-based strategies can actually vindicate the rationality of counterfactual imagination. Let's unpack.

5 | LAPLACIAN IMAGING IS RATIONAL

The feature of imaging we will resort to has been originally spotted, as far as we know, by Leitgeb and Pettigrew (2010), Pettigrew (2016), but never applied directly to the puzzle of imaginative use. To begin with, we need to broaden a bit our perspective on imaging.

The original Lewisian definition of imaging relied on Stalnaker's assumption that for each world and A there is a single closest A-world. If one doesn't like this assumption, which as is well known has been criticized by Lewis (1973) himself and others, one can generalize imaging to a framework with *set-selection functions*. A set-selection function, f, outputs, for each world w and A, the set $f_A(w)$ of most similar A-worlds, where the set may include more than one world. In imaging on A, then, one distributes probabilities among the worlds in such a set. Characterizations of generalized imaging have been proposed by Lewis (1981) himself, Gärdenfors (1982),Leitgeb (2017), and others.

A generalized transfer function, G, demands that all probabilities be moved to worlds in $f_A(w)$ (and that these add up to 1):

(1) If $w_1 \notin f_A(w)$, then $G_A(w, w_1) = 0$
(2) For each w: $\sum_{w_1 \in f_A(w)} G_A(w, w_1) = 1$

(A notationally similar presentation is in Pettigrew (2016), 203ff). A specific kind of generalized imaging with $f_A(w) = |A|$ has been called *Laplacian* by Joyce (2010). It pivots on the idea that the redistribution of probabilities should be uniform among the closest A-worlds. The intuitive idea: in the spirit of imaging, these worlds should be privileged, or get a probabilistic bonus, because they are the most similar among those which make A true. But looking inside this bunch, there's no reason to privilege any of *them* over any other.[2]

Let '$||A||$' stand for the cardinality of truth set $|A|$, i.e., for the number of worlds where A is true. Then the Laplacian transfer function L goes thus:

$$
L_A(w_m, w_n) = \begin{cases} 1 \text{ if } w_m, w_n \in |A| \text{ and } w_m = w_n \\ 0 \text{ if } w_m, w_n \in |A| \text{ and } w_m \neq w_n \\ \frac{1}{||A||} \text{if } w_m \notin |A| \text{ and } w_n \in |A| \\ 0 \text{ if } w_n \notin |A| \end{cases}
$$

The characteristic clause of Laplacian imaging is the third: the probability mass of a world w_m where A is not true is transferred and split in equal proportion among the relevant worlds w_n where A is true.[3]

---

[2] I mention that there's a change in view on this between Joyce (1999) and Joyce (2010). The former has Laplacian imaging as outlined above. The latter thinks one should take into account the differences in prior probabilities assigned to the worlds in $f_A(w)$: this 'Bayesianized' generalized imaging takes it closer to conditionalization.

[3] A very helpful referee remarks that this may give counterintuitive results in the assessment of some counterfactuals. Their example is: take a sequence of n independent tosses of a coin heavily biased to heads, whose outcome is known with certainty. Two natural ways to arrange worlds by relevant similarity are (i) worlds are close to the extent that the coin lands the same ways at them, (ii) worlds are all equally close. In case (i), say that by an amazing coincidence the coin always lands tails at the base world w. The counterfactual [1] 'If the coin had landed heads at least once, it would have landed heads only once' comes out certain, which is implausible given the coin is biased to heads. In case (ii), say that

Now when imaging works according to the Laplacian function (and so if $w_1 \in |A|$, then $P(w_1|_L A) = \sum_{w \in W} P(w)L_A(w, w_1)$; otherwise $P(w_1|_L A) = 0$), the following can be proved (Leitgeb and Pettigrew (2010); Pettigrew (2016) 201-5): *it minimizes expected inaccuracy among all probability distributions that give probability 1 to A, when inaccuracy is measured via the Brier score.*[4]

This is the crucial claim. But it's also a mouthful, so it requires some explanation. What is the Brier score? What does it mean that it measures the expected inaccuracy of a probability distribution representing degrees of belief?

The Brier score (see e.g. Section 10.2 of Titelbaum (2022) for an introduction) is named after statistician Glenn W. Brier, who proposed it in the Fifties as a way to evaluate the accuracy of weather forecasts. It is also called the *mean squared error* or *quadratic loss function*. It assesses the inaccuracy of a probability distribution. Inaccuracy, for such a distribution and so for a system of degrees of belief which is represented by it, is, intuitively, just its distance from the truth.

More precisely: the Brier score assesses inaccuracy by measuring in a certain way the distance of a probability distribution from the *ideal*, or maximally accurate, probability distribution. The ideal probability distribution is the one which gets things exactly right: it assigns probability 1 to all and only truths, 0 to all and only falsehoods. One may think of it, intuitively, as representing the beliefs of a maximally opinionated and omniscient, or God-like, agent: one who fully believes all and only the truths, and fully disbelieves all and only the falsehoods.

For a simple illustration of how it works, say you only have two claims to deal with: *A = Arif is in the kitchen*, and *B = There's no Beer in the fridge*. Say you are neutral on the former: your subjective probability or degree of confidence or of belief in the Arif-claim *A* is 0.5; you are moderately confident in the latter, giving the Beer-claim *B* a 0.7. Also, say that, as it happens, *A* and *B* are both true. So the ideal probability distribution here is the one assigning probability 1 to both *A* and *B*.

Now the Brier score measures your inaccuracy, as distance from the ideal probability distribution, as follows: it starts by (1) taking the difference between your degrees of belief in each of *A* and *B* and the ideal distribution separately (measures that work this way are therefore called *separable*): for the former, the difference is $|1 - 0.5| = 0.5$; for the latter, it's $|1 - 0.7| = 0.3$. Next, (2) it squares such a difference (that's why the rule is also labeled as 'squared' or 'quadratic'); and (3) it adds the results together: $0.5^2 + 0.3^2 = 0.25 + 0.09 = 0.34$.

A small complication: for our purposes, we need to relativize the rule to possible worlds. For there's a different possible world where, say, Arif is still in the kitchen but there's a lot of beer in the fridge, so *A* is still true but *B* is false. In such a situation, your beliefs are a bit less accurate, i.e., a bit further from the ideal. For the distance from the truth of your degree of belief in *B* is a bit greater now, namely $|0 - 0.7| = 0.7$; and so overall your inaccuracy is $0.5^2 + 0.7^2 = 0.25 + 0.49 = 0.74$. There are, of course, also worlds where Arif is not in the kitchen and there's a lot of beer in the fridge, etc., and we need to take all such variation across possible worlds into account.

---

the coin always lands heads at the base world. The counterfactual [2] 'If the coin had landed tails at least once, it would have landed tails at least twice' comes out almost certain, which is implausible, too. I have no idea on what exactly to say on this, but my first reaction would be to stick to measure (i): those worlds should not be all at the same distance from each other as per (ii). They should count as closer/more remote from each other, depending on the proportion of tosses on which the coin lands the same way, as per (i). But then it's not the fault of (i) that [1] comes out certain. The issue is that what happens at the base world w was unlikely to begin with: the coin always lands tails in spite of being heavily biased to heads, which is an amazing coincidence. So I think we have an 'implausibility in/implausibility out' situation. Measure (i) only ensures that the nearest worlds to a world where the unlikely happens are nearly as unlikely. This is in itself not that implausible, methinks.

[4] Precisely, it is proved that what minimizes expected inaccuracy by the Brier score is $P(-|_L A)$, where $P(X|_L A) = P(X \wedge A) + \frac{||X \wedge A||}{||A||}(1 - P(A))$: see Pettigrew (2016), 204-5 and 219-20.

Generalizing: given a finite set of claims $\Sigma = \{A_1, A_2, \ldots, A_n\}$, the Brier score, Br, gives the inaccuracy of probability distribution P at world w, $Br(P, w)$, as distance from the ideal probability distribution at w, I_w, thus:

$$Br(P, w) = (I_w(A_1) - P(A_1))^2 + (I_w(A_2) - P(A_2))^2 + \cdots + (I_w(A_n) - P(A_n))^2$$

That is, Br (1) takes the difference between $P(A_i)$ and $I_w(A_i)$ for each claim $A_i \in \Sigma$, (2) squares it, and (3) adds up the results.

The Brier score is popular among Bayesians as a measure of inaccuracy for credences or degrees of belief, and possibly the most popular, because it satisfies a set of conditions on scoring rules many find desirable (see again Titelbaum (2022), 339, for an introduction, and Part I of Pettigrew (2016) for a full-scale defense).

But why does (expected) accuracy matter? The research programme of *accuracy-first epistemology* (Joyce, 1998; David, 2001; Goldman, 2002a; Schoenfield, 2015; Pettigrew, 2016; Staffel, 2017) focuses on the idea that accuracy, that is, closeness to the truth, is the fundamental epistemic virtue of credences or (degrees of) belief. I have no new arguments for this claim besides the various ones presented in the existing literature. But I have a plain point regarding the extension of the view to imagination, of the kind we are investigating. After all, accuracy can be taken as a key epistemic aim of various representational mental states – i.e., states whose core function is to represent how things are in reality. In an exercise of counterfactual imagination, we assess some B, not plainly, but hypothetically, i.e., under the counterfactual supposition of A; but such an assessment can be more or less accurate, just as the plain assessment of B can be. In counterfactual, reality-oriented imagination we represent things, not as they actually are, but as they would be if something was the case. However, our aim is to get accurate beliefs on what *is* the case: when we wonder what would be or have been the case if we jumped the river, if the atom was ionized, if the driver had not been distracted, etc., we look through the hypothetical scenario and back into the real world, where our epistemic interests lie. To the extent that our degree of belief that we would succeed in getting to the other side of the river, under the counterfactual supposition that we jump, is accurate, our belief in the counterfactual that if we jumped over the river, we would succeed in getting to the other side, is accurate about the *actual* world. To the extent that our degree of belief that nobody else would have killed Kennedy, under the counterfactual supposition that Oswald had not killed him, is accurate, our belief in the counterfactual that if Oswald had not killed Kennedy, then nobody else would have, is accurate about the *actual* world.

So we now have a precise solution to the Puzzle of Epistemic Use – if one accepts the assumptions I have unpacked so far. One assumption is that the relevant sort of imagination is adequately understood as a form of simulated belief revision, governed by a principle of minimal alteration. This, as I claimed in section 1, seems to be a view on which there is already some consensus in the philosophy of imagination.

Another assumption is that the minimal alteration at issue, for counterfactual imagination, can be characterized via imaging. The idea that, counterfactually supposing that A, we minimally hypothetically revise our beliefs as prescribed by imaging on A, as argued in section 4, gives a natural way to capture the insight that in reality-oriented imagination we look at the most similar scenarios where the initial supposition A is true, to check the likelyhood of the relevant Bs there.

Another assumption is that the Brier score is a (or the) good measure of the inaccuracy of our degrees of belief. A final one is that accuracy, or closeness to the truth, is a (or the) key epistemic virtue of representational mental states in general, and therefore, as claimed in this section, in particular of counterfactual imagination.

Then one gets that the latter, as governed by (Laplacian generalized) imaging, maximizes such a virtue by minimizing expected inaccuracy. Under such assumptions, we have a vindication of the rationality of counterfactual imagination: not only is it minimally altering in a formally precise sense of the term, but also, it is expected to be maximally accurate as a hypothetical update of our belief system in the light of counterfactual suppositions.

6 | IMAGING VS CONDITIONALIZATION?

But how does this square with the fact that the very same Brier score rules that planning to *conditionalize* on *A* minimizes expected inaccuracy? (As proved by Greaves and Wallace, 2006). Is Bayesian conditionalization in trouble?

Conditionalization and imaging can be reconciled by considering that belief revision (actual or hypothetical) can be carried out in two very different ways, following a well-established distinction from the AI literature on formal theories of belief revision. This is the distinction between revisions triggered by newly received information on an unchanged environment, for which the go-to account is the AGM theory of Alchourrón, Gärdenfors, and Makinson (Alchourrón et al., 1985); and revisions (sometimes called 'updates') triggered by changes in the worldly environment itself, for which the go-to account is often the KGM theory of Katsuno, Grahne, and Mendelzon (Grahne, 1991; Katsuno and Mendelzon, 1992). Leitgeb and Segerberg (2005) and Leitgeb (2017) conjecture that if AGM-style static qualitative belief revision has conditionalization as its natural probabilistic counterpart, so can KGM-style belief change in a dynamic environment have imaging as its natural probabilistic counterpart. (The recent Eva et al. (2022) investigates the proportion *AGM : conditionalization = KGM : imaging* in formal detail.)

So the corresponding explanation of what is going on in the probabilistic case, is that accuracy-based arguments for conditionalization à la Greaves and Wallace (2006) are, in a sense, static: they only vindicate one's *planning* to conditionalize on *A*, should one receive the information that *A* (Titelbaum, 2022, 366-7). But counterfactual imagination is more dynamic: when we imagine counterfactually that Oswald has not killed Kennedy, we suppose that things *have gone otherwise* than they actually have, not just that we received new information about an unchanged world. In this sense, imaging and conditionalization may coexist in peace as governing two distinct, rational suppositional procedures.

ORCID

Francesco Berto https://orcid.org/0000-0003-3246-657X

REFERENCES

Alchourrón, C., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic*, *50*, 510–30.

Arcangeli, M. (2019). *Supposition and the Imaginative Realm*. Routledge, new york.

Arlo-Costa, H., & Parikh, R. (2005). Conditional probability and defeasible inference. *Journal of Philosophical Logic*, *34*, 97–119.

Badura, C. (2021). How imagination can justify. In Badura, C. and Kind, E., editors, *Epistemic Uses of Imagination*. Routledge, New York.

Baltag, A., & Smets, S. (2008). Probabilistic dynamic belief revision. *Synthese*, *165*, 143–66.

Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford University Press, Oxford.

Berto, F. (2022). *Topics of Thought: The Logic of Knowledge, Belief, Imagination*. Oxford University Press, Oxford.

Byrne, R. (2005). *The Rational Imagination. How People Create Alternatives to Reality*. MIT Press, Cambridge, MA.

Currie, G., & Ravenscroft, I. (2002). *Recreative Minds*. Oxford University Press, Oxford.

David, M. (2001). Truth as the epistemic goal. In Steup, M., editor, *Knowledge, Truth, and Duty*. Oxford University Press, Oxford.

Davies, J. (2019). *Imagination. The Science of Your Mind's Greatest Power*. Pegasus Books, New York.

Eva, B., Shear, T., & Fitelson, B. (2022). Four approaches to supposition. *Ergo*, *8*, 58–98.

Evans, J., & Over, D. (2004). *If*. Oxford University Press, Oxford.

Gärdenfors, P. (1982). Imaging and conditionalization. *Journal of Philosophy*, *79*, 747–60.

Gillies, A. (2004). Epistemic conditionals and conditional epistemics. *Noûs*, *38*, 585–616.

Goldman, A. (2002a). *Pathways to Knowledge: Private and Public*. Oxford University Press, New York.

Goldman, A. (2002b). *Simulating Minds*. Oxford University Press, Oxford.

Grahne, G. (1991). Updates and counterfactuals. In Allen, J., Fikes, R., and Sandewall, E., editors, *Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann, San Mateo.

Greaves, H., & Wallace, D. (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, *115*, 607–32.

Gregory, D. (2016). Imagination and mental imagery. In Kind, A., editor, *The Routledge Handbook of Philosophy of Imagination*, pages 97–110. Routledge, New York.

Hajek, D. (1989). Probabilities of conditionals. *Journal of Philosophical Logic*, *18*, 423–8.

Jackson, F. (1979). On assertion and indicative conditionals. *The Philosophical Review*, *88*, 565–89.

Johnson, M., & Raye, C. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.

Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, *65*, 575–603.

Joyce, J. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press, Cambridge.

Joyce, J. (2010). Causal reasoning and backtracking. *Philosophical Studies*, *147*, 139–54.

Katsuno, H., & Mendelzon, A. (1992). Belief revision. In Gärdenfors, P., editor, *On the Difference Between Updating a Knowledge Base and Revising It*. Cambridge University Press, Cambridge.

Kind, A. (2001). Putting the image back in imagination. *Philosophy and Phenomenological Research*, *62*, 85–109.

Kind, A. (2016). Imagining under constraints. In Kind, A. and Kung, P., editors, *Knowledge Through Imagination*, pages 145–59. Oxford University Press, Oxford.

Kind, A., & Kung, P., editors, (2016). *Knowledge through Imagination*. Oxford University Press, Oxford.

Kosslyn, S., & Moulton, S. (2009). Mental imagery and implicit memory. In Markman, K., Klein, W., and Suhr, J., editors, *Handbook of Imagination and Mental Simulation*, pages 35–52. Taylor and Francis, New York.

Langland-Hassan, P. (2016). On choosing what to imagine. In Kind, A. and Kung, P., editors, *Knowledge Through Imagination*, pages 61–84. Oxford University Press, Oxford.

Leitgeb, H. (2017). Imaging all the people. *Episteme*, *14*, 463–79.

Leitgeb, H., & Pettigrew, R. (2010). An objective justification of bayesianism ii: The consequences of minimizing inaccuracy. *Philosophy of Science*, *77*, 236–72.

Leitgeb, H., & Segerberg, K. (2005). Dynamic doxastic logic: Why, how, and where to? *Synthese*, *155*, 167–90.

Lewis, D. (1973). *Counterfactuals*. Blackwell, Oxford.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, *95*, 581–9.

Lewis, D. (1981). Causal decision theory. *Australasian Journal of Philosophy*, *59*, 5–30.

Lewis, D. (1999). Why conditionalize? In Eagle, A., editor, *Philosophy of Probability: Contemporary Readings*, pages 403–7. Routledge, New York.

Mahtani, A. (2012). Diachronic dutch book arguments. *Philosophical Review*, *121*, 443–50.

Nichols, S., & Stich, S. (2003). *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press, Oxford.

Pearl, J. (2013). *Causality. Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford University Press, Oxford.

Ramsey, F. (1990). General propositions and causality. In Mellor, D., editor, *Philosophical Papers*, pages 145–63. Cambridge University Press, Cambridge.

Schoenfield, M. (2015). Bridging rationality and accuracy. *Journal of Philosophy*, *112*, 633–57.

Staffel, J. (2017). Accuracy for believers. *Episteme*, *14*, 39–48.

Stalnaker, R. (1968). A theory of conditionals. In Rescher, N., editor, *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, pages 98–112. Blackwell, Oxford.

Titelbaum, M. (2022). *Fundamentals of Bayesian Epistemology 2: Arguments, Challenges, alternatives*. Oxford University Press, Oxford.

Van Fraassen, B. (1995). Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic*, *24*, 249–77.

Van Leeuwen, N. (2013). The meanings of imagines part i: Constructive imagination. *Philosophy Compass*, *8*, 220–30.

Williamson, T. (2016). Knowing by imagining. In Kind, A. and Kung, P., editors, *Knowledge Through Imagination*, pages 113–23. Oxford University Press, Oxford.

Williamson, T. (2020). *Suppose and Tell. The Semantics and Heuristics of Conditionals*. Oxford University Press, Oxford.