

Is it the Principal Principle that implies the Principle of Indifference?

(preprint, to appear in G. Hofer-Szabó, L. Wroński (eds.), *Making it formally explicit: Probability, Causality, and Determinism*, Springer, hopefully 2017)

Balázs Gyenis¹ and Leszek Wroński²

¹Institute of Philosophy, Research Center for the Humanities, Hungarian Academy of Sciences

²Institute of Philosophy, Jagiellonian University

January 20, 2017

Abstract

Hawthorne et al. (2015) argue that the Principal Principle implies a version of the Principle of Indifference. We show that what the Authors take to be the Principle of Indifference can be obtained without invoking anything which would seem to be related to the Principal Principle. In the Appendix we also discuss several Conditions proposed in the same paper.

The argument of Hawthorne et al. (2015) (“Authors” from here on) that the Principal Principle implies what they take to be the Principle of Indifference is based on their Proposition 2, which (using their notation, and remembering that X “says that the chance at time t of proposition A is x and E is any proposition that is compatible with X and admissible at time t ,” *ibid.* p. 1.¹) can be restated as follows:

Proposition 2. Let $P(A|X) = P(A|XE) = P(A|FXE) = P(A|(A \leftrightarrow F)XE)$. Then from $x \neq 0$, $x \neq 1$ and from the Principal Principle $P(A|X) = x$ it follows that $P(F|XE) = 1/2$.

Unfortunately it is unclear what the essence of Proposition 2 has to do with the Principal Principle. One can easily show that its “indifference” message is simply a consequence of imposing a sufficient number of independence constraints on an arbitrary probability space:

Proposition 2’. Let $P(A|X) = P(A|XE) = P(A|FXE) = P(A|(A \leftrightarrow F)XE)$. Then from $P(A|X) \neq 0$, $P(A|X) \neq 1$ it follows that $P(F|XE) = 1/2$.

¹We note here that the Authors require also that F be “contingent and atomic”. From their footnote 1 it is clear that by this they do not mean that F should be an atom of the considered algebra of propositions, and rather that it should not be “logically complex”. We believe that the mathematics of the current paper, and specifically the Counterexample in the Appendix is acceptable from that point of view.

The proof is a straightforward calculation which does not make use of the Principal Principle. As it can be readily seen by comparing the Authors' Proposition 2 with Proposition 2' the only tenuous connection between the Authors' Proposition 2 and the Principal Principle comes from that in order to conclude that $P(F|EX) = 1/2$ one needs to rule out the extreme cases when $P(A|X) = 0$ or $P(A|X) = 1$; Proposition 2 achieves this by setting $x \neq 0$, $x \neq 1$ and applying the Principal Principle, but this is an overkill, since $P(A|X) \neq 0$, $P(A|X) \neq 1$ could be directly achieved by simply requiring that learning event X does not make A impossible or certain.

Thus the Principal Principle does not have much to do with the essence of Proposition 2. The Authors themselves also make a side-remark on p. 4. that "interestingly, [their] line of argument does not depend on the structure of the proposition X ," but they do not follow up on it and they do not seem to come to the realization that the indifference message of Proposition 2 is a mere consequence of the independence constraints: their Superprincipal principle, with which the Authors generalize the Principal Principle, also makes assumptions about the structure of proposition X , and their generalized argument given subsequently still relies on equating a conditional probability $P(A|XE)$ with a number $x \neq 0$, $x \neq 1$ to reach the conclusion. Assuming that $P(A|XE) = x$ and $x \neq 0$, $x \neq 1$ is still an overkill to merely ensure $P(A|XE) \neq 0$, $P(A|XE) \neq 1$, which as we can see in Proposition 2' is all that is needed to arrive at $P(F|XE) = 1/2$. And thus the Superprincipal principle does not have much to do with the essence of Proposition 2 either.

Notice that our claim is not that it is possible to obtain the Principle of Indifference from something weaker than the Principal Principle: this would even strengthen the original point the Authors wanted to make! We say rather that the argument does not seem to have any connection to the Principal Principle, since it is an inference from a set of independence assumptions to a conclusion about independence, not relying on any statement regarding (credences about) chances.

The Authors devote some time to "defending" their Condition 2 which is intended to furnish one of the premises of their Proposition 2. It can be restated as follows:

Condition 2. If $P(A|XE) = P(A|X)$, $P(F) = P(F|XE)$, and the Principal Principle $P(A|X) = x$ holds, then $P(A|XE(A \leftrightarrow F)) = P(A|X)$.

The Authors state that Condition 2 "must hold because [it] encapsulate[s] core intuitions about defeat" (p. 4.); however, one can show with a counterexample that Condition 2 is not a theorem of probability theory. Instead of searching for ways to "defend" Condition 2 it might be instructive to ask what additional supposition (S) the Authors would need to add to the premises of Condition 2 to make it a theorem. It is easy to find candidates for (S) that express probabilistic requirements in terms of A, E, F, X (and possibly other events) without referring to the value x ; however if (S) took such form, then the same (S) could also be added to the premises of

Condition 2'. If $P(A|XE) = P(A|X)$, $P(F) = P(F|XE)$ holds, then $P(A|XE(A \leftrightarrow F)) = P(A|X)$

to make this latter a theorem. This is since if (S) does not refer to x , then the additional assumption $P(A|X) = x$, which is what differentiates Condition 2 from Condition 2', will

clearly not be used in the derivation. Thus if the Authors succeed in finding and defending such an (S) with which Condition 2 furnishes the desired premise of Proposition 2, then with the same (S) Condition 2' also furnishes the corresponding premise of Proposition 2'. Since neither Condition 2' nor Proposition 2' invokes the Principal Principle, we again would complete an argument for what the Authors' interpret as the Principle of Indifference without invoking the Principal Principle. (For the mentioned counterexample, proof of Proposition 2', and additional comments on the Authors' arguments regarding their Condition 1 and Condition 2 the Reader is referred to the Appendix.)

Acknowledgements. Balázs Gyenis was supported by the National Research, Development and Innovation Office, K 115593 grant. Leszek Wroński was supported by Tomasz Placek's MISTRZ grant "Probability, modalities and Bell theorems" of the Foundation for Polish Science. The work on the current paper was carried out as a part of the "Probability, Causality and Determinism" Bilateral Mobility Grant of the Hungarian and Polish Academies of Sciences, NM-104/2014.

The authors would like to thank Gergei Bana and Christian Wallmann for insightful comments on a draft of this paper.

Appendix

Proof of Proposition 2'

Before the proof for later purposes we slightly reformulate Proposition 2':

Proposition 2'. Let (\mathcal{L}_X, P_X) be a probability space, $A, E, F \in \mathcal{L}_X$, and let $P_X(A) = P_X(A|E) = P_X(A|FE) = P_X(A|(A \leftrightarrow F)E)$. Then from $P_X(A) \neq 0$, $P_X(A) \neq 1$ it follows that $P_X(F|E) = 1/2$.

Proof. Using the following identification for the probabilities for different conjunctions:

$P_X(\cdot)$	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8
A					x	x	x	x
E			x	x			x	x
F		x		x		x		x

(so that $AEF = o_8$, $P_X(AEF) = p_8$ etc.) we can restate the conditions as

$$p_8/(p_4 + p_8) = p_8/(p_3 + p_8) \tag{1}$$

$$p_8/(p_4 + p_8) = p_5 + p_6 + p_7 + p_8 \tag{2}$$

$$p_8/(p_3 + p_8) = (p_7 + p_8)/(p_3 + p_4 + p_7 + p_8); \tag{3}$$

we also require $p_i \geq 0$, $\sum_{i=1}^8 p_i = 1$, and $p_3 + p_4 + p_7 + p_8 > 0$, $p_4 + p_8 > 0$, $p_3 + p_8 > 0$ for the conditional probabilities to make sense.

By (1) we get that either $p_3 = p_4$ or $p_8 = 0$. If $p_8 = 0$ then from (2) $P_X(A) = p_5 + p_6 + p_7 + p_8 = 0$ and all other conditions can be satisfied. If $p_8 \neq 0$ then assuming $p_3 = p_4$ from

(3) we get that either $p_7 = p_8$ or $p_4 = 0$. If $p_4 = 0$ then from (2) $P_X(A) = p_5 + p_6 + p_7 + p_8 = 1$ and all other conditions can be satisfied. Assuming $p_8 \neq 0$, $p_4 \neq 0$ and thus that $p_3 = p_4$ and $p_7 = p_8$, $P_X(F|E) = (p_4 + p_8)/(p_3 + p_4 + p_7 + p_8) = (p_4 + p_8)/(2p_4 + 2p_8) = 1/2$ and all other conditions can be satisfied. To sum up, either $P_X(A) = 0$, or $P_X(A) = 1$, or $P_X(F|E) = 1/2$.
QED

Remarks on Conditions 1 and 2 and a counterexample

The premises of the Authors' Proposition 2 are supposed to be furnished by the following two Conditions:

“Condition 1. If E is not a defeater and XE contains no information that renders F relevant to A , then EF is not a defeater.”

“Condition 2. If E is not a defeater and XE contains no information relevant to F , then $E(A \leftrightarrow F)$ is not a defeater.” (ibid., p. 2)

which are, as mentioned earlier, to be understood in a context where X says that the chance at time t of proposition A is x . The Authors call E a “non-defeater” (w.r.t. A) if $P(A|XE) = P(A|X) = x$ holds.

Unfortunately the Authors “take the supposition that XE contains no information that renders F relevant to A to imply that $P(A|FXE) = P(A|XE)$ ” (ibid., p. 2). In other words, according to the Authors, if XE contains no information that renders F relevant to A then XE screens off F from A . This is surely too strong. The most natural mathematical counterpart of the Author's natural language expression in Condition 1 seems to be the one corresponding closely to “if F is not relevant to A then XE does not make it relevant”, namely

$$P(A|F) = P(A) \rightarrow P(A|FXE) = P(A|XE). \quad (4)$$

With that in place, Condition 1 becomes formally the following:

Condition 1 (reformulated). If $P(A|XE) = P(A|X)$, $P(A|F) = P(A) \rightarrow P(A|FXE) = P(A|XE)$, and the Principal Principle $P(A|X) = x$ holds, then $P(A|FXE) = P(A|X)$.

Formula (4) is indeed weaker than the Authors' screening off condition, as evidenced by the following Counterexample (which—we claim—falsifies also the original Conditions 1 and 2):

Counterexample 1 Let \mathcal{L} be the Boolean algebra generated by the logically independent propositions A, E, F ,² and $X = “Ch(A) = 2/5”$ and let P be the probability on \mathcal{L} that assigns $1/10$ to the propositions $\neg X \neg A \neg E \neg F$, $\neg X \neg A E \neg F$, $X \neg A \neg E \neg F$, $X \neg A E \neg F$, and

²And so F can be assumed not to be “logically complex”, as the Authors wish.

assigns $1/20$ to the rest of the algebraic atoms. (\mathcal{L}, P) then has the following properties:

$$\frac{2}{5} = P(A|XE) = P(A|X) = \frac{2}{5} \quad (5)$$

$$\frac{1}{2} = P(A|F) \neq P(A) = \frac{2}{5} \quad (6)$$

$$\frac{1}{2} = P(A|FXE) \neq P(A|XE) = \frac{2}{5} \quad (7)$$

$$\frac{1}{2} = P(A|FXE) \neq P(A|X) = \frac{2}{5} \quad (8)$$

$$\frac{2}{5} = P(F) = P(F|XE) = \frac{2}{5} \quad (9)$$

$$\frac{1}{3} = P(A|XE(A \leftrightarrow F)) \neq P(A|X) = \frac{2}{5} \quad (10)$$

(i) Due to (5),(6), and (9) (\mathcal{L}, P) is a counterexample to

$$\begin{aligned} P(A|XE) &= P(A|X) = x \\ P(A|F) &= P(A) \rightarrow P(A|FXE) = P(A|XE) \end{aligned}$$

$$P(A|FXE) = P(A|X) = x,$$

that is, to the reformulated Condition 1 (see below)

(ii) Due to (5),(6),(7), and (8) (\mathcal{L}, P) is a counterexample to Condition 1. (On the assumption that from the fact that due to (6) and (7) we see that conditioning on XE leaves both the probability of A and the probability of F intact we can infer that “ XE contains no information that renders F relevant to A ”.)

(iii) Due to (5),(9), and (10) (\mathcal{L}, P) is a counterexample to Condition 2.

The interpretation of the propositions involved has to matter since the Authors only provide a single intuition boosting example to illustrate the reasonableness of Condition 1, and the plausibility of their illustration rests on a proper reading of what it is for a characterization of a situation to not contain information that renders two events relevant to each other. Apart from this sole example the Authors do not provide any philosophical analysis of the plausibility of Condition 1, because they claim, with their Proposition 1, that Condition 1 “provably holds.” However, it holds only if “containing no information that renders F relevant to A ” is expressed as “screening off F from A ”; our Counterexample also shows that it does not hold if we follow the formulation of the condition more accurately, for example by using (4).

There are of course other formal candidates for “ XE containing no information that renders F relevant to A ” which could be considered, of various logical strength. One which perhaps could be promising is $P(A|F) - P(A) = P(A|FXE) - P(A|XE)$. However, if we use *that* in Condition 1, it agains falls prey to our Counterexample 1—as evidenced by the clause (ii). The situation is similar with all other candidates we have explored.

Since their Proposition 1 is no proof of Condition 1 as reconstructed above, the Authors would need to provide a philosophical justification for it. But since the justification for Condition 2 also depends on the availability of justification of Condition 1—the Authors counter one of the main objections against Condition 2, namely that it fails under a definition of admissibility put forward by Meacham (2010), by claiming that that Meacham’s admissibility is too restrictive is evidenced by the fact that it entails a violation of the supposedly provable Condition 1—the premises of the main Proposition 2 stay insufficiently motivated, which then carries over to whatever conclusion one can draw from it.

To wager philosophical bets on the potential success of justification of Condition 1 and Condition 2 we remark that our Counterexample also evidences a probability space in which both of these conditions fail. Universal quantification (or at least the existence of a sound argument for typicality) over events and spaces matters since the mere fact that some events in some probability spaces have a probability $1/2$ would not establish a connection between the Principal Principle and the Principle of Indifference.

We emphasize that our remarks regarding the appropriate formal reconstruction of the Authors’ verbal formulation of Condition 1 is independent from the argument developed in the main text.

As for some further remarks regarding Condition 2, we note that the flavor “under some conditions, since we are ignorant of whether an independence constraint holds, we assume that it does” it possesses is similar to “under some conditions, since we are ignorant of the probability, we assume that it is $1/2$ ”—which is the Principle of Indifference. This raises the suspicion that we are proving what we already put in; indeed a requirement of probabilistic independence is of comparable strength as narrowing down the probability to a single value, as we are now going to illustrate.

Thus let us consider now the Authors’ inference from $P(F|XE) = 1/2$ to $P(F) = 1/2$ right after their Proposition 2, which is the basis for their philosophical claim that the Principle of Indifference should also hold for the unconditional prior. It should be clear that this inference is only valid if Condition 2 indeed holds, but the assumptions of their Proposition 2 can hold without Condition 2 being true (or, alternatively, assumptions of Proposition 2’ can hold without Condition 2’ being true). One can also quantify how frequently this is the case. Let us assume that (\mathcal{L}_X, P_X) satisfies the conditions of our Proposition 2’ and hence $P_X(F|E) = 1/2$. Using the terminology of the proof of Proposition 2’, let \mathcal{L} be generated by adding an additional algebraic atom o_0 to \mathcal{L}_X , let A, E, F be the same as in the proof and let $X = \{o_1, \dots, o_8\}$; then with the probability $P(Y) = p_0 \cdot \chi_Y(o_0) + (1-p_0) \cdot P_X(Y|X)$ the new (\mathcal{L}, P) also satisfies the conditions of Proposition 2’ and $P(Y|X) = P_X(Y)$. It is then easy to show that $P(F) = 1/2$ only when $p_0 = 1 - 1/(2(p_2 + p_4 + p_6 + p_8))$. Thus even though in all of the so-defined extensions $P(F|EX) = 1/2$, there is only one in which $P(F) = 1/2$ (and thus in which the independence constraint $P(F) = P(F|XE)$ holds) while in all of the uncountably many others $P(F) \neq 1/2$. Loosely speaking it appears, then, that in this case we should expect Condition 2 to hold only as frequently as we would expect $P(F) = 1/2$ to hold if we set its value randomly.

References

- Hawthorne, J., Landes, J., Wallmann, C. and Williamson, J. (2015). The Principal Principle Implies the Principle of Indifference, *The British Journal for the Philosophy of Science*,
URL: <http://bjps.oxfordjournals.org/lookup/doi/10.1093/bjps/axv030>
- Meacham, C. J. (2010). Two Mistakes Regarding the Principal Principle, *British Journal for the Philosophy of Science* **61**: 407–431.