Philosophy of Science (2025), **00**, 1–21
doi:10.1017/psa.2024.40

ARTICLE

Deep Uncertainty and Incommensurability: General Cautions about Precaution

Rush T. Stewart

Department of Philosophy, University of Rochester
Email: rush.stewart@rochester.edu

(Received 07 June 2023; revised 08 November 2023; accepted 21 August 2024)

Abstract

The precautionary principle is invoked in a number of important personal and policy-related decision contexts. Peterson shows that certain ways of making the principle precise are inconsistent with other criteria of decision making. Some object that the results do not apply to cases of deep uncertainty or value incommensurability, which are alleged to be in the principle's wheelhouse. First, I show that Peterson's impossibility results can be generalized considerably to cover cases of both deep uncertainty and incommensurability. Second, I contrast an alternative way of giving voice to the precautionary impulse.

1. Introduction

An important tradition in decision theory has worked to advance the maxim "Live to fight another day" as a principle of rational choice. The resulting decision rules prioritize avoiding worst-case outcomes. A prominent manifestation of this effort is the precautionary principle.[1] The precautionary principle is routinely discussed in connection with consequential social decisions made under conditions of significant uncertainty. Such conditions emerge in contexts involving law Steele (2006), environmental policy (United Nations, 1992; Sprenger, 2012), and health policy (Wingspread, 1998), including, of considerable recent interest, the appropriate responses to pandemics (Kamran, 2020). Very roughly, the principle counsels decision makers, at least under certain circumstances, to be driven primarily by avoiding potential catastrophic outcomes even if accepting the risk of catastrophe comes with the possibility of substantial benefits. Given widespread appeals to the precautionary principle in high-stakes policy decisions, understanding the virtues and vices of the principle and its consistency with other important principles of rational choice is of first-order importance.

In contrast to many presentations couched in somewhat vague language, Peterson (2006) presents a few attempts to formalize the principle in more precise terms.

[1] Other manifestations, I take it, include conservative choice rules like maximin, Γ-maximin, the Hurwicz criterion (for certain settings), and minimax regret.

On the basis of some mathematical observations, Peterson argues that the precautionary principle, once it is made suitably precise, is *incoherent* as a decision rule. Essentially, he shows that the formulations of the principle that he considers are inconsistent with other putative norms of rational choice. According to some, Peterson has shown "convincingly that these decision rules conflict with attractive principles of rational choice" (Sprenger, 2012, p. 883).

Others, however, find nothing in Peterson's analysis that tells so decisively against the precautionary principle. For example, Boyer-Kassem argues that "Peterson's argument fails to establish the incoherence of the Precautionary Principle" in large part because of allegedly overly restrictive assumptions about uncertainty and preference (2017a, p. 2026). The precautionary principle's primary applications, it is claimed, arise under less restrictive assumptions. In many important policy contexts, we are forced to act under massive uncertainty and without a determinate assessment of the value of certain potential outcomes. Estimating the knock-on effects and long-term consequences of our actions and policies is extremely challenging.[2] For example, will philanthropic aid for schools and clinics in poor countries free up funds for military use by despotic regimes (Deaton, 2013, ch. 7)? Will decelerating artificial intelligence research help stave off human extinction or delay significant life improvement for millions (Bengio et al., 2023)? According to expected utility theory, decision makers should choose options with the greatest probability-weighted average utility. In the presence of severe uncertainty, some claim, maximizing expected utility—the dominant normative approach to decision making—is infeasible. Decision makers may be unable or unwilling to make definitive comparisons of likelihood, let alone to assign numerically precise probabilities. In other cases, making determinate comparisons of value may be infeasible. How, precisely, does the value of saving the life of an 80-year-old compare to the value of saving the life of a 10-year-old (Caplan, 2021)? In the realm of private life, how does one trade off patriotic and pacifist commitments when they come into conflict (Dewey and Tufts, 1932)? Not only might we lack the sort of introspective access to preferences and values that might permit such evaluations, but we may also lack the determinate preferences and values themselves. If the precautionary principle has interesting applications under severe uncertainty or value incommensurability, but Peterson's assumptions rule such cases out, then the principle may yet have important roles to play.

In this article, I generalize Peterson's assumptions about uncertainty and value commensurability. I relax the requirement that the comparative likelihood must be complete. But the assumptions about uncertainty may not be where the real action is. In his reply to Boyer-Kassem, Peterson says, "The second [objection], about value commensurability, is arguably his most important concern" (Peterson, 2017, p. 2036). He goes on to sketch a possible reply, but it does not convince Boyer-Kassem: "When answering the second part of my objection, Peterson suggests an escape route: change

[2] For a dramatic description of uncertainty confronting policymakers and everyday decision makers alike, consider the problem of *cluelessness* (Lenman, 2000). One need not be a thoroughgoing consequentialist to concede that the consequences of an action are extremely important to its moral evaluation. Even seemingly morally obvious decisions can have unintended and fraught consequences: Will saving a drowning child in the Danube allow that child to grow up to commit mass atrocities (Mogensen, 2021)?

the scope of the theorem so that it applies to incommensurable outcomes. I doubt this can be done—not only does the Archimedean condition need a reformulation, but also the total order condition" (Boyer-Kassem, 2017b, p. 2040). Here, I show that this can, in fact, be done. To address Boyer-Kassem's concern about the possibility of incommensurability in preference, I relax totality and weaken the Archimedean condition significantly in a few different ways. We need not even assume that desirability assessments are given by a binary relation. We can take choice functions as primitive. Even dropping some of Peterson's other assumptions entirely, impossibilities remain. This extension is important because objectors to the significance of Peterson's results claim that because of the restrictive assumptions involved, the precautionary principle is confined "to a fraction of the cases discussed in the literature," and even if the precautionary principle were "indeed incoherent as Peterson claims, it would only be proven with a small scope and would actually not apply to the most interesting cases" (Boyer-Kassem, 2017a, p. 2029). Although they do not forestall every possible complaint, of course, the results that follow help us to see that appeals to incommensurability will not necessarily skirt the spirit of Peterson's critique. Recording the observations in sections 3 and 5 hopefully helps to focus disputes on the most relevant issues. In section 6, I explain how these generalizations respond to three criticisms of Peterson's results in the literature. In the final part of the article, I consider a different way of thinking about precaution in decision making. This alternative framing allows for both severe uncertainty and incommensurability and locates roles for both trade-off and precautionary reasoning to play.

2. Preliminaries

Let S be a non-empty, finite set of states with typical element s. These states may be provisional, not very specific, subject to revision, and so forth. Let O be a non-empty set of outcome or consequence elements represented by lowercase Latin letters (except s). Alternatives or options are functions $X : S \to O$ that associate states with outcomes. So $X(s) \in O$ is the outcome of act X in state s. Let $\mathcal{A}$ denote the set of all options or alternatives. Outcomes themselves can be embedded in the set of options by the usual technique of identifying an outcome x with constant alternatives $\mathfrak{c}_x$: for all $s \in S, \mathfrak{c}_x(s) = x$.

Let Σ be an algebra of events over S. Because S is finite, we can just take $\Sigma = 2^S$. Let $\dot{\succsim} \subseteq \Sigma \times \Sigma$ be a binary relation on Σ, which we will interpret as giving qualitative probability comparisons. The expression $E \mathrel{\dot{\succsim}} F$ means that the event E is at least as probable as the event F. Given the appeals to pseudo-rationalizability and imprecise probabilities (IP) that follows, it is natural to assume that $\dot{\succsim}$ is a "partial likelihood relation" because such relations admit "multiprior" or IP representations.[3] That is, if $\dot{\succsim}$ is a partial likelihood relation, there exists a set $\mathbb{P}$ of probability functions on (S, Σ) such that, for all $E, F \in \Sigma$, $E \mathrel{\dot{\succsim}} F$ if and only if $P(E) \geq P(F)$ for all $P \in \mathbb{P}$. Boyer-Kassem alleges that "Peterson's treatment of uncertainties lacks generality" (2017a, p. 2026). Partial likelihood relations are rather general and, unlike what Peterson assumes, allow that some events cannot be compared in terms of their

[3] In particular, it is natural to assume that $\dot{\succsim}$ satisfies the properties discussed in the literature on partial likelihood relations (e.g., Harrison-Trainor et al., 2016). Because such properties will play no substantive role here, I omit discussion of them.

likelihoods. A limiting case worth noting is the one in which $\dot{\succsim}$ is representable by the set of *all* probability functions on (S, Σ). It is plausible, then, that all types of *probabilistic* uncertainty of concern in the literature on the precautionary principle can be seen as special cases of partial likelihood relations.[4]

A *menu* is a non-empty, finite subset of $\mathcal{A}$. Menus represent decision problems. A *choice function* is a set-valued function $C : 2^{\mathcal{A}} \backslash \{\emptyset\} \to 2^{\mathcal{A}} \backslash \{\emptyset\}$ such that $C(\mathcal{S}) \subseteq \mathcal{S}$ and $C(\mathcal{S}) \neq \emptyset$ for any menu $\mathcal{S} \subseteq \mathcal{A}$.[5] Choice functions can be thought of as selecting the acceptable/admissible/choiceworthy options in a menu. Selecting the best alternatives according to some complete and transitive preference relation generates a particular choice function, but there are choice functions that cannot be reduced to binary comparisons. In general abstract choice theory, certain properties of choice functions are especially interesting and play important roles. Perhaps the most central such property is Sen's property α.

$$\mathcal{S} \subseteq \mathcal{T} \Rightarrow \mathcal{S} \cap C(\mathcal{T}) \subseteq C(\mathcal{S}) \tag{α}$$

Property α—also known as *contraction consistency*, *heritage*, and *Chernoff*—requires that acceptable options remain acceptable when other items are removed from the menu. Sen's property β requires that if two options, X and Y, are both acceptable in a menu, and X remains acceptable when the menu is expanded to include additional options, then Y must also remain acceptable.

$$\mathcal{S} \subseteq \mathcal{T}, X, Y \in C(\mathcal{S}), \text{ and } X \in C(\mathcal{T}) \Rightarrow Y \in C(\mathcal{T}) \tag{β}$$

Together, properties α and β characterize those choice functions that are rationalizable by a weak order (complete and transitive) preference relation; that is, C satisfies α and β if and only if there exists a complete and transitive relation $\succsim \subseteq \mathcal{A} \times \mathcal{A}$ such that, for all menus $\mathcal{S} \subseteq \mathcal{A}$,

$$C(\mathcal{S}) = \{X \in \mathcal{S} : X \succsim Y \text{ for all } Y \in \mathcal{S}\}.$$

In other words, such choice functions can be regarded as choosing the best elements from a menu according to the relation $\succsim$. In the presence of property α, property γ is strictly weaker than property β.

[4] Even though Peterson appeals only to comparative judgments of likelihood, Boyer-Kassem and others claim that this is too restrictive because it "requires that one comes up with a set of all possible outcomes for the decisions under consideration" and that "a most important case in which [the precautionary principle] applies is when outcomes are poorly defined" (2017a, p. 2029). I suppose my treatment of uncertainty will not be wholly satisfactory to some advocates of the precautionary principle. But I do not find it plausible that the precautionary principle has any interesting application without *some* specification of relevant outcomes, even if "poorly defined." As stated earlier, we can allow that the possible outcomes as we deal with them are provisional, not very specific, subject to revision, or poorly specified in some sense; nothing in the mathematics of the results requires more. Moreover, Boyer-Kassem apparently concedes that Peterson's handling of uncertainty can be defended by less general means than those pursued here (2017b, p. 1).

[5] In general, it is plausible to associate contexts of greater uncertainty or indeterminacy in preference with *less* stringent judgments of admissibility; fewer options can be excluded from the choice set. Non-emptiness is a natural assumption if we interpret menus as the set of *all* options in a decision problem. Some decision theorists deal with concerns about this assumption by stipulating that an *abstain* or *status quo* option is always available (e.g., Fishburn, 1973). There is also work that explores permitting empty choice sets (e.g., Aizerman, 1985).

$$C(\mathcal{S}) \cap C(\mathcal{T}) \subseteq C(\mathcal{S} \cup \mathcal{T}) \qquad (\gamma)$$

Options that are acceptable in both menu $\mathcal{S}$ and menu $\mathcal{T}$ are acceptable in the menu $\mathcal{S} \cup \mathcal{T}$. Together, property α and property γ characterize binary choice. A binary choice function is one for which there exists *some* binary relation—not necessarily a weak order—that rationalizes it.

Following a common convention in the more formal literature on the precautionary principle (e.g., Peterson, 2006; Boyer-Kassem, 2017a; Stefánsson, 2019), we use letters $p, q, \ldots$ to denote catastrophic or fatal outcomes in O (except for lowercase letters corresponding to the designation of an act: so x, x_i, etc. are not necessarily catastrophic when representing generic outcomes of an act X). I will not offer any substantive account of such outcomes here but will just flag that the ability to draw such a line in a substantive way makes strong measurability assumptions.[6]

3. A generalization of Peterson's first impossibility theorem

Peterson states impossibility results for two versions of the precautionary principle. I will state and prove generalizations of each. Peterson's first impossibility result relies on a formulation of the precautionary principle given by his *PP*(α). Informally, "If one act is more likely to give rise to a fatal outcome than another, then the latter should be preferred to the former; and if the two acts are equally likely to give rise to a fatal outcome, then they should be equi-preferred" (2006, p. 597). In contrast to Peterson's *PP*(α), here, *PP*$(\alpha)_c$ is formulated for choice functions rather than binary relations, let alone a particular type of binary relation like a weak order. Say that some outcome x is *not more choiceworthy* than another outcome y when $\mathfrak{c}_y$ is acceptable in a choice between it and $\mathfrak{c}_x$.

> Let X be an alternative such that there is at least one outcome that is not more choice worthy than p. Then,
>
> (1) if the likelihood of an outcome that is not more choice-worthy than p is greater for X than for Y, X is not acceptable in the menu *{X, Y}*; $\qquad$ (PP$(\alpha)_c$)
>
> (2) if an outcome that is not more choiceworthy than p is as likely for X as for Y, then neither X nor Y is ruled out.

Informally, PP$(\alpha)_c$ says that if one act is more likely to lead to a catastrophic outcome than another, then the former is unacceptable. If the two acts are equally likely to yield a catastrophic outcome, then neither is uniquely acceptable in a binary choice between them. (More formal statements of the various assumptions included in the theorems are provided in the Appendix.) Even if proponents of precautionary reasoning intend a stronger principle that legislates choice even when the sorts of likelihood comparisons made in PP$(\alpha)_c$ are unavailable, it seems plausible to think such a principle would imply something very much like PP$(\alpha)_c$ in those special cases

[6] Given the ability to draw such a line, it is plausible to require that, for a menu of constant acts, if some noncatastrophic constant option is available, a catastrophic option is never chosen.

in which such likelihood comparisons are available.[7] Our assumptions allow that some (other) outcomes may not be comparable in terms of likelihood and that some acts (including X and Y themselves in the statement of $PP(\alpha)_c$ may not be comparable in terms of a binary preference relation.

Next, Cov_c is a choice-theoretic analogue of Peterson's *covariance* condition intended to capture the idea that "everything else being equal, the less likely a bad outcome is to occur, the better" (2006, p. 597).

> Let x_i, x_j be possible outcomes of alternative X such that $\mathfrak{c}_{x_j}$ is uniquely acceptable in a binary choice between it and $\mathfrak{c}_{x_i}$, and X is more likely to result in x_i. Let X' be the alternative obtained from X by swapping which states yield x_i and x_j. Then, in a binary choice between X and X', alternative X' is uniquely acceptable. $(\mathrm{Cov_c})$

Effectively, alternative X' is just like X, only more likely to yield x_j than x_i. Because $\mathfrak{c}_{x_j}$ is uniquely acceptable in the choice between it and $\mathfrak{c}_{x_i}$, according to Cov_c, X' is uniquely acceptable in a binary choice between it and the original alternative X.

We can now state a rather substantial generalization of Peterson's first impossibility result.

Theorem 1. *$PP(\alpha)_c$ and Cov_c are inconsistent.*

It is worth stressing that theorem 1 requires neither the dominance nor the ordering assumptions of Peterson's impossibility result for his analogue of $PP(\alpha)_c$ (2006, Theorem 1, p. 600). So, in addition to generalizing Peterson's $PP(\alpha)$ and covariance conditions, Theorem 1 generalizes Peterson's result by dropping two further assumptions completely. Among other things, this helps us to isolate the source of inconsistency.

4. Incommensurability

In general, talk of (unique) admissibility cannot be construed as coded talk of an alternative being *better* than others in a menu. The latter description assumes a binary relation that determines choiceworthiness. But only in special cases—namely, when properties α and β are both satisfied—does choiceworthiness reduce to binary comparisons.

Peterson makes the standard assumption that preferences are given by a weak order.[8] This implies that any two alternatives X and Y can be compared, with X at

[7] Perhaps not the second clause. A proponent of precautionary reasoning may appeal to other considerations in cases like the one in clause 2. There is no analogue of this clause in the second formulation of the precautionary principle considered later in the article.

[8] Although Peterson explicitly states that preferences *totally* order acts, in the appendix, he works with weak preference relations and allows for indifference. His informal version of $PP(\alpha)_c$ and the proof of his theorem 1 explicitly appeal to indifference. If he means that *strict* preference forms a total order, there can be no nontrivial indifference because total order strict preferences are semi-connex and antisymmetric (he says "asymmetric"). If he means that *weak* preference is a total order, then there can be no indifference because of antisymmetry (asymmetry is inconsistent with the totality of weak order preferences). So I assume that by "total order," Peterson means to refer to what is often called a "weak order."

least as good as Y or Y at least as good as X. Neither $\text{PP}(\alpha)_c$, 2, nor, consequently, theorem 1 appeals to a weak order. In particular, the impossibility reported there does not rely on any general commensurability assumptions. The theorem does not even assume that the choice function satisfies the central internal coherence constraint property α, only that C is, in fact, a choice *function*. So the prospect of incommensurability providing a context for coherent application of the precautionary principle as formulated in $\text{PP}(\alpha)_c$ seems dim by theorem 1. Nevertheless, incommensurability is a central issue in debates about the precautionary principle, and it is important to explain how relaxing the assumption of a weak order desirability ranking in generalizing Peterson's impossibility results allows for forms of incommensurability. I will illustrate this with the property of *path independence*, although I stress that the generalizations do not assume even this much.

In order to respond to Boyer-Kassem's criticism—which we return to in section 5—that many and some of the most important applications of the precautionary principle are to cases lacking determinate desirability comparisons, cases in which certain outcomes are *incommensurable*, Peterson considers extending the scope of his observations so that it applies to incommensurable options. He then suggests a choice rule for handling such options: "If X and Y are incomparable, and the agent is rationally permitted to choose X, then it is also rationally permitted to choose Y" (Peterson, 2017, p. 2037, with options capitalized for consistent notation). Peterson concedes that "[t]his new condition is controversial (because some think it may make us vulnerable to money pump arguments) [. . .] I am not claiming that this is the correct analysis of incomparability, but the example suggests that if we knew how incomparable values should be linked to rational choice it would also be possible to construct a corresponding impossibility theorem" (Peterson, 2017, p. 2037). It is about this point that Boyer-Kassem expresses skepticism that the program sketched can be carried out in detail because various assumptions, including the Archimedean and weak order preference conditions, require modification.

That the precautionary principle is inconsistent with some choice rule or other is not by itself a challenge to the precautionary principle. For a plausible challenge, it is important either to identify a plausible decision rule for incommensurability or to show that no choice rule at all supports the coherent articulation of precaution along the lines Peterson pursues. The results reported in sections 3 and 5 pursue the second strategy. In this section, however, I want to consider not the rule that Peterson tentatively suggests but path-independent choice. This is a natural and nontrivial generalization of the standard assumption of a weak order that has been used in thinking about incommensurability, as I explain later in the article.[9] Of course, *adding* the assumption of path independence to my generalizations of Peterson's impossibility results for the precautionary principle will not help skirt the limitations reported there, but my point is that the leaner observations recorded here apply to

[9] In addition to its more mathematical interest (Danilov and Koshevoy, 2005), path independence finds interesting application in the contexts of social choice (Plott, 1973), matching theory (Chambers and Yenmez, 2017), decision theory for imprecise probabilities (Levi, 1980), formal epistemology (Rott, 2001), and population ethics (Stewart, MS). In the paper in which Plott introduced the property, for example, he notes that, compared to the assumption of weak order social preference, the assumption of a merely path independent social choice function opens up certain possibilities in the context of social choice theory.

path-independent choice functions, among others. As a result, simply appealing to incommensurability—which can be naturally modeled for path-independent choice functions—is no automatic out.

Intuitively, path independence says that the order of consideration or presentation of options in a menu does not affect the final choice.

$$C(\mathcal{S} \cup \mathcal{T}) = C(C(\mathcal{S}) \cup C(\mathcal{T})) \tag{PI}$$

In a sense, path-independent choice functions allow harder choice problems ($S \cup T$) to be decomposed into easier problems (S and T). That choice from a menu is independent of the way options are divided up for consideration is an important feature of standard weak order rationalizability that path independence retains (cf. Danilov and Koshevoy, 2005). A consequence of path independence, then, is that certain forms of manipulation are excluded. A decision maker's choice from a menu cannot be manipulated by presenting options in a different order.[10] Aizerman and Malishevski show that such choice functions admit an interesting representation. If C is a path-independent/pseudo-rationalizable choice function, then there is a *set* $\{\succsim_i\}_{i\in I}$ of weak orders such that, for any menu $\mathcal{S}$,

$$C(\mathcal{S}) = \bigcup_{i\in I} max_S \succsim_i. \tag{1}$$

In other words, rather than selecting the optimal elements in a menu with respect to a single preference relation, a path-independent choice function selects those elements that are optimal according to at least one preference relation in the set $\{\succsim_i\}_{i\in I}$. One interpretation of the Aizerman and Malishevski decomposition is that there is indeterminacy in preference when $|I| > 1$. For example, it could be the case that, according to one permissible way of evaluating things, $X \succ_i Y$, and according to another, $Y \succ_j X$.[11]

So, if C is path independent, then there exists a set of weak orders $\{\succsim_i\}_{i\in I}$ such that, for any menu $\mathcal{S}$, $C(\mathcal{S}) = \bigcup_{i\in I} max_S \succsim_i$. In words, C selects those alternatives in $\mathcal{S}$ that are maximal according to *some* relation $\succsim_i$ in the set $\{\succsim_i\}_{i\in I}$. The elements of $\{\succsim_i\}_{i\in I}$ can be interpreted as rival but permissible assessments of the alternatives in terms of desirability. Such sets arise in cases of indeterminacy or vagueness in desirability assessments. And at least in some such cases, the rankings may correspond to various permissible ways of trading off certain fundamental valuations of the alternatives,

[10] Another way to think about path independence is as the conjunction of two basic "coherence" properties. Moulin proves that a choice function satisfies PI if and only if it satisfies both Property _ and Aizerman's Axiom (Aiz) (1985, Lemma 6).

$$C(\mathcal{T}) \subseteq \mathcal{S} \subseteq \mathcal{T} \Rightarrow C(\mathcal{S}) \subseteq C(\mathcal{T}) \tag{Aiz}$$

(Aiz) According to Aiz, impermissible options do not become permissible by removing some other impermissible options from the menu. Choice functions satisfying Property α and Aiz have also been called pseudorationalizable (e.g., Aizerman and Malishevski, 1981).

[11] Inspection of the proof of the equivalence of PI and the conjunction of α and Aiz reveals that the equivalence holds generally and does not depend on the assumption that A is finite (Aizerman and Malishevski, 1981; Moulin, 1985, Lemma 6). The decomposition of a path independence choice function into maximizing each weak order in a particular set (Aizerman and Malishevski, 1981, Theorem 3; Moulin, 1985, Theorem 5) is stated for finite A, though see (Pedersen, 2009).

certain ways of compromising between different values. There are a number of ways to define *categorical* desirability from a set of desirability orderings. The simplest may be to just take the intersection of the orderings: $\succsim = \bigcap_{i \in I} \succsim_i$ (e.g., Sen, 2004, p. 672). This proposal allows for the possibility that $X \succ Y$ even if $X \succsim_i Y$ for some $i \in I$. That is, it is possible that strict categorical desirability comparisons hold that are not unanimously shared among the $\succsim_i$. Another proposal is for categorical desirability to consist of the unanimously held strict desirability comparisons and the unanimously held indifferences (cf. Sen, 2004, p. 674). Levi proposes a more complicated account according to which categorical desirability consists of unanimous strict desirability, unanimous indifference, *and* unanimous weak desirability (2008). In all three of these proposals, if $X \succ_i Y$ and $Y \succ_j X$ for some $i, j \in I$, then X and Y are categorically incommensurable.

Like Peterson's rule, path independence is phrased in choice-theoretic rather than relation-theoretic terms—although, as we have seen, there is a set-based relation-theoretic representation via pseudo-rationalizability. Path independence, however, does *not* vindicate Peterson's rule for incommensurable options.

Example 1. *Let $\mathcal{A} = \{X, Y, Z\}$. Let C be a path-independent choice function on $\mathcal{A}$ that is rationalized by $\{\succsim_1, \succsim_2\}$, defined as follows:*

$$X \succ_1 Y \succ_1 Z$$

$$Z \succ_2 Y \succ_2 X$$

Here, and Y are categorically incommensurable. Relative to $\succ_1$, alternative X is more desirable; according to $\succ_2$, the opposite comparison obtains. Peterson's suggested rule would imply that Y is acceptable in the menu $\{X, Y, Z\}$ because X is. But this is not the case for C: $C(\{X, Y, Z\}) = \{X, Z\}$. The alternative Y is not maximal according to any permissible evaluation of the options in that menu.

The fundamental idea is that $X \in C(\{X, Y\})$ does not imply that, for some categorical desirability weak order $\succsim$, $X \succsim Y$. It could be that X is at least as desirable as Y, but it could also be the case that X is strictly more desirable than Y, or even—and this is the key point—that X and Y are incommensurable. In example 1, X and Y are categorically incomparable; it is indeterminate whether X is more desirable than Y or Y is more desirable than X.

Path independence is one important way of thinking about incommensurability. I will return briefly to it in section 7. My point here is not to argue that path independence is definitely mandatory for rational choice but to illustrate how theorems 1 and 2—which assume only a choice function and not path independence—extend the limitations for the precautionary principle to forms of incommensurability. Arguably, any plausible theory of decision making is committed to the bare assumption of choice function.

5. A generalization of Peterson's second impossibility result

Peterson's second theorem is the more "refined and powerful one," according to Boyer–Kassem (2017a, p. 2028). The formalization of the precautionary principle in this theorem is the weakest and most general that Peterson considers, "so weak that it cannot reasonably be refuted by any advocate of the precautionary principle" (2017,

p. 599). Still, the theorem is not fatal for the precautionary principle, Boyer-Kassem claims, because it assumes general commensurability in desirability. And he finds Peterson's subsequent reply to that worry unconvincing: "When answering the second part of my objection, Peterson suggests an escape route: change the scope of the theorem so that it applies to incommensurable outcomes. I doubt this can be done—not only does the Archimedean condition need a reformulation, but also the total order condition" (Boyer-Kassem 2017b, p. 2040). I have explained how incommensurability is consistent with the extremely minimal assumption of a choice function. My task now is to generalize the relevant precautionary principle and Archimedean condition accordingly.

Peterson's informal statement of his weakest formulation of the precautionary principle is this: "If one act is more likely to give rise to a fatal outcome than another, then the latter should be preferred to the former, given that (i) both fatal outcomes are equally undesirable, and (ii) not negligibly unlikely, and (iii) the nonpreferred act is sufficiently more likely to lead to a fatal outcome than the preferred one" (2006, p. 599). In his formalization of this property, Peterson assumes that the relevant fatal/catastrophic outcome for both acts is equi-preferred to the best fatal outcome (namely, p) (2006, p. 601). Clause (ii) implies that we are outside the context of the potential application of the de minimis principle, according to which sufficiently improbable outcomes can be ignored or treated very differently in deliberation (Peterson, 2002; Lundgren and Stefánsson, 2020). The choice-theoretic version replaces assumptions about preference and desirability with assumptions about acceptability (and again, we do not assume completeness of the partial likelihood relation).

> Let Y be an alternative that is more likely to result in a catastrophic outcome than alternative X. Y is not acceptable in the menu $\{X,Y\}$ if
>
> (1) there is exactly one outcome of each of X and Y—x_i and y_j, respectively—that is catastrophic, and neither catastrophic outcome is more choice worthy than the other; $(\mathrm{PP}(\delta)_c)$
> (2) neither x_i nor y_j is negligibly unlikely, and
> (3) Y is sufficiently more likely to lead to y_j than X is to lead to x_i.

$\mathrm{PP}(\delta)_c$ generalizes Peterson's constraint because it does not require that both catastrophic outcomes are *equally* undesirable, although the property is implied by that special case. It requires instead that neither catastrophic outcome is more choiceworthy than the other. $\mathrm{PP}(\delta)_c$ only governs cases of acts that have a single catastrophic outcome. This is suggested by Peterson's more formal articulation of this version of the principle in his appendix. As with $\mathrm{PP}(\alpha)_c$, this special case is plausibly implied by any stronger precautionary principle that governs additional cases, such as when acts have multiple possible catastrophic outcomes. Unlike Peterson's version, $\mathrm{PP}(\delta)_c$ does not assume that the relevant catastrophic outcomes are indifferent to p, although again, clause 1 is implied by that special case as well. We also do not assume, as Peterson does, that X is strictly preferred to Y, only that Y is not acceptable in the binary choice between it and X.

The only other assumption of Peterson's that is needed to state a choice-theoretic generalization of his second impossibility theorem is a version of what he calls an *Archimedean* condition. In spirit, Peterson's Archimedean condition is similar to the Archimedean or continuity assumptions associated with von Neumann and Morgenstern expected utility theory. It captures a sense in which trade-offs should be considered in choice. As Peterson puts it, "advocates of the precautionary principle must be willing to admit that, to some extent, both the likelihood and the desirability of an outcome matter" (2006, p. 599). His informal statement of this assumption is as follows: "If the relative likelihood of a nonfatal outcome is increased in relation to a strictly better nonfatal outcome, then there is some (nonnegligible) decrease of the relative likelihood of a fatal outcome that counterbalances this precisely" (2006, p. 599). As I explain in section 6, it is the assumption of a trade-off Archimedean condition on which many objectors focus. My informal statement of the choice-theoretic version of this condition again generalizes assumptions about preference or desirability to choiceworthiness and tracks Peterson's formal version.

> There are atleast two alternatives $X, Y \in \mathcal{A}$ such that Y is acceptable in the menu $\{X, Y\}$, but
>
> (1) X and Y have the same set of possible outcomes,
> (2) no outcome for either alternative is negligibly unlikely, (Arch$_c$)
> (3) there is exactly one catastrophic outcome x_i, and
> (4) Y is sufficiently more likely than X to lead to x_i.

Arch$_c$ is considerably weaker than Peterson's Archimedean condition on a few counts. First, by using choice functions, we do not require indifference between the acts X and Y, only that Y is acceptable in the menu $\{X, Y\}$. This latter requirement is consistent with indifference between X and Y, with Y being strictly preferred to X, and also with X and Y being incommensurable according to a categorical desirability relation.[12] Second, Peterson's condition quantifies over all alternatives. By contrast, *Arch$_c$* only asserts the existence of a pair of options in $\mathcal{A}$ meeting the stated assumptions.[13] Third, *Arch$_c$* does not commit to a particular means—such as how many likelihood relations between outcomes are modified—by which Y attains its admissibility in the choice between it and X. Arch$_c$ is so weak that it does not really resemble an Archimedean condition, that label being retained for continuity with the literature on Peterson's results. The interest in working with a much weaker condition is not limited to the fact that the associated mathematical results are stronger; the main point is that the assumption is more difficult to deny. If a proponent of precaution wishes to deny *Arch$_c$*, the scope of the precautionary principle is more significantly restricted still. Its applicability demands that not even two alternatives as are mentioned in *Arch$_c$* exist,

[12] Stefánsson's "Weak Archimedes" also weakens Peterson's assumption, but by replacing indifference with weak preference for X (2019, p. 1219). So, Arch$_c$ weakens Stefánsson's Weak Archimedes, too.

[13] Compare the way in which Sen weakens his liberalism condition in proving a stronger version of the impossibility of a Paretian liberal Paretian liberal (1970).

that not even this much trade-off reasoning is allowed. The existence of such a pair of alternatives is, of course, consistent with the incommensurability of the options X and Y and with extremely widespread incommensurability in general.

We can now state the generalization of Peterson's second impossibility theorem.

Theorem 2 *$PP(\delta)_c$ and $Arch_c$ are inconsistent.*

Boyer-Kassem suggests that there is already a conflict between Peterson's most general statement of the precautionary principle and his Archimedean condition without invoking the additional assumptions of Peterson's theorem (2017a, p. 2031). Theorem 2 verifies that this is true even when those assumptions are stated choice-theoretically. As with theorem 1, theorem 2 generalizes Peterson's corresponding result, not only by substantially weakening the precautionary and Archimedean assumptions but also by dropping two other assumptions—those of a weak order desirability ranking and a certain dominance condition—altogether.

6. Responses to some reservations about Peterson's results

About Peterson's impossibility theorems, Sprenger writes, "The source of the problem is the intuition that both the probability and desirability/potential harm of an outcome matter and that they can, to some extent, be traded off against each other. This view is deeply entrenched in most accounts of rational decision making" (2012, pp. 883–884). It is the Archimedean condition that encodes some of this deeply entrenched intuition. The Arch_c formulation of Peterson's Archimedean condition and theorem 2 allow us to offer responses to three objections that have been voiced in the literature. Although I will not argue that Peterson is definitely correct on all of the issues involved, I think the results presented in this article show that some objections focus on inessential features of Peterson's observations and thereby fail to be satisfactory responses to some of the concerns that they raise.

First, Boyer-Kassem objects to the commensurability of catastrophic and noncatastrophic outcomes assumed by Peterson's original formulation of the Archimedean condition. He writes, "if one accepts PP (in Peterson's sense), one is committed to this view of incommensurability between fatal and nonfatal outcomes. Now, the problem is that the Archimedean condition is saying exactly the opposite: by stating that a change in the likelihood of nonfatal outcomes can be compensated by a change in the likelihood of fatal outcomes, it assumes that the desirability of fatal and nonfatal outcomes can be compared—even if one change of likelihood has to be much smaller than the other—and thus that fatal and nonfatal outcomes are commensurable" (2017a, p. 2031). The source of the conflict, Boyer-Kassem seems to be suggesting, is the commensurability between certain types of outcomes—namely, fatal and nonfatal—that is illicitly assumed by the Archimedean condition. As it is stated here, however, Arch_c is formulated in terms of a choice function rather than a binary desirability relation. Given our (lack of) assumptions about C, we cannot infer commensurability from acceptability. Furthermore, Arch_c only makes an assumption about a single pair of alternatives. So, *even if* Arch_c were making an assumption about the commensurability of fatal and nonfatal outcomes, the scope of the commensurability assumed would be very minimal.

These same points regarding Arch_c can be used to address a second, related objection that Boyer-Kassem raises: "My second criticism against the Archimedean condition is that it assumes a value commensurability between outcomes in general [. . .] the Archimedean condition assumes that all outcomes can be compared, so that changes in the likelihood of some outcomes can be compensated by changes in the likelihood of some other outcomes. This gives another reason to reject the Archimedean condition" (2017a, p. 2031). Here, the worry concerns the assumption of commensurability "in general" rather than just between fatal and nonfatal outcomes. But to repeat, Arch_c does not assume general value commensurability, certainly not that all outcomes can be compared. For one thing, the assumptions in place on choice functions in theorem 2 do not secure comparability of alternatives (or all constant alternatives/outcomes) by a binary desirability relation. For another, Arch_c makes a claim only about a single pair of alternatives.

A third objection to interpreting Peterson's results as trouble for the precautionary principle comes from Steel (2015). Steel's objection, like Boyer-Kassem's, focuses on the Archimedean condition: "For what increase of credibility of catastrophe relative to poor would *precisely* offset the advantage accruing from the increase of the credibility of excellent relative to good? I submit that there is no non-arbitrary way to answer such a question" (2015, p. 42). Steel is referring here to Peterson's informal statement of his Archimedean condition. The "precise offsetting" shows up in Peterson's more formal statement of the condition as a claim about indifference between two alternatives: an initial alternative and one that results from it by increasing the likelihoods of both a catastrophic outcome and an excellent outcome. But Arch_c and t 2 help us to see that the focus on "precisely offsetting" is something of a red herring because, as the theorem establishes, an assumption of indifference is inessential to the derivation of a contradiction.

I do not anticipate that the approach presented here will have left Peterson's critics and advocates of the precautionary principle more generally bereft of replies. I am not even attempting to reply to all criticisms that have been voiced about Peterson's interpretation of his results, focusing in this section on some concerns raised about the Archimedean condition. Perhaps some of the foregoing objections can be repurposed to articulate new objections to the choice-theoretic assumptions presented here, even if, as they're stated, they fall short. Alternatively, one might object to the foregoing formulations of the precautionary principle on different grounds.[14] I hope, however, that this approach can help to structure further debate and clarify which issues are really at stake.

7. An alternative approach to the precautionary principle

In essence, the impossibility we encounter in theorem 2 is a conflict between trade-off reasoning, in the form of Arch_c, on the one hand, and precautionary reasoning, in the form of $\mathrm{PP}(\delta)_c$, on the other. One possibility worth exploring is that the problem

[14] There are a number of other interpretations of the precautionary principle that view it as something other than a formal principle of decision theory (see Resnik, 2021, ch. 4.8 and references therein). Proponents of these alternative interpretations may regard Peterson's critique as "a proof against a straw man" (2021, p. 83, fn. 12). I take no stance here on the viability of such interpretations or on the relevance of my results to them.

arises because both types of reasoning are applied simultaneously. A different approach, which I will now consider, would be to apply them lexicographically.[15]

The most developed account of weighing trade-offs is, of course, expected utility theory. Standard expected utility theory, however, might—and has been thought to—fall short when it comes to deep uncertainty and value incommensurability. Certainly, many fans of the precautionary principle think so. States of uncertainty are restricted to numerically determinate probability judgments. And a basic starting assumption in the classical formulations of expected utility is that desirability weakly orders the alternatives so that there is no incommensurability. Generalizations of expected utility theory have been developed that drop both of the evidently restrictive assumptions of completeness and numerically determinate probabilities. A natural and very general extension of the standard theory allows for *sets* of probabilities and *sets* of utilities rather than the assumptions of a single probability function and a single utility function for a decision maker.[16] Before, in motivating partial likelihood relations, I mentioned how two events may fail to be comparable in terms of likelihood. Similar remarks apply when we come to the expected utilities of options. Assume for the moment that desirability is a weak order and has a utility representation. If $\mathbb{P}$ is a set of probabilities representing $\dot{\succsim}$, it could be the case that, relative to one $P \in \mathbb{P}$, $EU_P(X) > EU_P(Y)$, whereas relative to another $P' \in \mathbb{P}$, $EU_{P'}(Y) > EU_{P'}(X)$. In such cases, one might think, and some decision theories imply, that considerations of expected utility secure no categorical preference between X and Y.

A natural generalization of the injunction to maximize expected utility when probabilities and utilities may not be determinate is *E-admissibility*, propounded prominently by Levi (e.g., 1980). Where $\mathbb{P}$ is a set of probabilities and $\mathbb{U}$ a set of utilities, the *E*-admissible options are those that maximize expected utility with respect to some $P \in \mathbb{P}$ and some $U \in \mathbb{U}$. Let's suppose that utility is determinate for ease of exposition. In Levi's interpretation, elements of $\mathbb{P}$ are permissible probability assessments, so *E*-admissibility restricts choice to those options that are best according to some permissible way of evaluating risk. Put differently, and abstracting from reference to states for simplicity, the *E*-admissible options are given by the set

$$\{X \in \mathcal{S} : \exists P \in \mathbb{P} \forall Y \in \mathcal{S} \;\; EU_P(X) \geq EU_P(Y)\}.$$

Note that a choice function induced by *E*-admissibility is pseudo-rationalizable, with the set of expected utility rankings as the rationalizing set of weak orders. *E*-admissibility reduces to expected utility maximization when $\mathbb{P}$ is a singleton (given our assumption that utility is determinate also). *E*-admissibility has been considered by some to be excessively permissive (for a recent example of such a critique, see, e.g., Mogensen and Thorstad, 2022). In response to this concern, Levi and others have considered a certain combination of *E*-admissibility with some other rule applied lexicographically as a second-tier criterion for narrowing the set of admissible

[15] As I explain later, Bartha and coauthors consider a distinct but related proposal, which I discovered while writing this article.

[16] That decision making with imprecise probabilities, and the strand of thinking deriving from Levi's work in particular, allows for the expression of some amount of precautionary reasoning has been emphasized previously (Sahlin, 2006).

options. The most prominent such two-tiered rule is especially interesting for my purposes in the present section, but presenting it requires a bit more setup.

In earlier work, Hansson suggests a model of precautionary reasoning differing from the ones we have been considering so far: "The maximin rule can be used as a formal version of the precautionary principle" (1997, p. 293). Each alternative is associated with a *security level*, that alternative's worst outcome for any possible state. Maximin selects those alternatives that have maximal security levels, thereby *maxi*mizing the *min*imum. As with other proposed formulations of the precautionary principle, the extent to which maximin ignores trade-offs has been the subject of criticism 1957, pp. 279-280). Some find that even as a formulation of the precautionary principle, the conditions of applicability of maximin are either too restrictive or that the rule disregards information outside of those conditions. Bartha and DesRoches think identifying the precautionary principle with maximin makes it too discontinuous with expected utility theory: "we reject the Maximin interpretation of PP because such an identification makes it impossible to clarify the relationship between PP and ordinary expected utility maximization. Maximin operates in the framework of decisions under ignorance; standard decision theory applies to decisions under risk. There is no element in common. [...] our goal is to show that PP is more closely related to standard decision theory than the Maximin interpretation allows" (2021, p. 8709). On its own, maximin may well be too conservative, wasteful of valuable information, subject to convincing counterexamples, and so forth. But, following an important strand of research in decision theory and *pace* Bartha and DesRoches, I want to consider a role for maximin in a unified setting that allows for both ignorance and risk.

In the setting of imprecise probabilities, the manifestation of the conservative, maximin approach to decision making is sometimes called Γ-maximin (Gärdenfors and Sahlin, 1982; Gilboa and Schmeidler, 1989). Suppose that $\mathbb{P}$ is a set of probability distributions on some common measurable space. According to Γ-maximin, we should restrict choice to those options with the greatest minimal expected utility. Continuing with the simplifying assumption that utility is determinate, the choice set is given by

$$\left\{X \in \mathcal{S} : \inf_{P \in \mathbb{P}} EU_P(X) \geq \inf_{P \in \mathbb{P}} EU_P(Y) \text{ for all } Y \in \mathcal{S}\right\}.$$

Under complete uncertainty—when no probability distributions are excluded from $\mathbb{P}$—maximin and Γ-maximin coincide (e.g., Berger, 1985, p. 216). Γ-maximin is a significant generalization, and it clarifies at least one way of understanding the relationship between the precautionary principle and ordinary expected utility maximization. When $\mathbb{P}$ is a singleton (and utility is determinate), Γ-maximin and expected utility maximization coincide. Perhaps Γ-maximin could serve as the sort of bridge Bartha and DesRoches are looking for, connecting maximin and (even substantial generalizations of) standard decision theory.[17]

[17] Various criticisms of Γ-maximin as a standalone decision rule have been voiced (e.g., Al-Najjar and Weinstein, 2009). (But see (Siniscalchi, 2009; Hill, 2020).) Both Seidenfeld (2004) and Troffaes (2007), for example, compare it unfavorably to Levi's E-admissibility. Adjudicating this debate is not my concern here.

I'll call the rule that first eliminates all options that are not E-admissible and then applies Γ-maximin to the surviving options $E+\Gamma$. This two-tiered or lexicographic rule has been studied in the literature on IP decision theory (e.g., Levi, 1986; Seidenfeld et al., 2012). As with Γ-maximin and E-admissibility, $E+\Gamma$ reduces to expected utility maximization when $\mathbb{P}$ is a singleton. To the extent that E-admissibility is a generalized form of trade-off reasoning and Γ-maximin captures some of the precautionary impulse, $E+\Gamma$ is one way of reconciling these general reasoning styles. Whether E-admissibility satisfies Arch_c is more or less a matter of the richness of the set of outcomes: provided two such alternatives as described in the four clauses exist, Y will be acceptable in a choice between it and X so long as it maximizes expected utility with respect to some $P \in \mathbb{P}$. In some cases, Γ-maxmin will satisfy $\text{PP}(\delta)_c$. The key clause is the third: Y is sufficiently more likely to lead to its catastrophic outcome than X is to lead to its catastrophic outcome. If the greater likelihood of Y's catastrophic outcome makes it such that X's minimal expected utility (across $P \in \mathbb{P}$) is greater than Y's minimal expected utility, then $\text{PP}(\delta)_c$ is satisfied: Y is not acceptable in a choice between it and X. To repeat, for $E+\Gamma$, both X and Y must maximize expected utility with respect to some probability assessment that has not been excluded (that is, some $P \in \mathbb{P}$).

Two further comments on Γ-maximin and $\text{PP}(\delta)_c$ are in order. One way of securing a tighter link between the two is to use Γ-maximin to flesh out the content of Y's being "sufficiently more likely" to lead to its possible catastrophic outcome than X is to lead to its possible catastrophic outcome. If we define "sufficiently more likely" here as *the likelihood of Y's catastrophic outcome is greater than that of X's and is such that X's minimal expected utility is greater than Y's minimal expected utility*, the satisfaction of $\text{PP}(\delta)_c$ is secured. In the context of $E+\Gamma$, $\text{PP}(\delta)_c$ would be secured only at the second tier, for the restricted application to E-admissible options. Alternatively, rather than appealing to Γ-maximin as a second-tier criterion, we could simply impose $\text{PP}(\delta)_c$ or a suitable strengthening of it as a tie-breaking rule after first restricting choice to E-admissible options.

In some recent publications, Bartha and coauthors explore an alternative lexicographic model of precautionary reasoning (2017; 2021; 2023). Their approach requires avoiding catastrophe *first* and *then* maximizing expected utility—a sort of reversal of the order of operations when compared to $E+\Gamma$. In favor of $E+\Gamma$, one might point out that choiceworthy options are forced to pass trade-off analysis with respect to at least *some* feasible probability assessment, which, as proponents and critics alike point out, can be quite a weak requirement in the presence of deep uncertainty. We might think of $E+\Gamma$ as making precautionary reasoning palatable to (broad-minded) expected utility partisans. Strong partisans of precautionary reasoning, on the other hand, might find Bartha et al.'s lexical approach more congenial to their initial inclinations by not subordinating precaution to trade-off reasoning but, instead, advancing the converse subordination. Others may find the lexicographic approach objectionable in general, reasoning that *vast advantages* in security are worth *some* sacrifice in trade-off superiority and, similarly, that *vast advantages* on the trade-off ledger surely license at least *some* sacrifice in security level.[18] Whether such a view can be precisely and coherently articulated remains to be seen, as far as I am aware.[19] At any rate, there are multiple routes for further exploring a reconciliation of trade-off and precautionary reasoning.

8. Conclusion

Peterson shows that some reasonable formulations of the precautionary principle are inconsistent with other plausible decision-theoretic principles. As stated, his results do not cover the cases of deep uncertainty and value incommensurability, where advocates of precautionary reasoning claim the precautionary principle has important roles to play. But as has been shown here, extensions of Peterson's results can be established for these contexts as well. The generalizations help us to see that certain criticisms of Peterson's results may be interpreted as objecting to inessential features of the tension he has identified. However, for those who endorse a commitment to something like orthodox trade-off reasoning, precautionary reasoning may yet have important applications. In contexts of deep uncertainty and value incommensurability, precautionary reasoning can be appealed to lexicographically—as in the case of the $E + \Gamma$ rule—to help prune the set of choiceworthy options.

Acknowledgments. Thanks to Jean Baccelli, Michael Nielsen, Mark Swails, and two anonymous referees for helpful feedback. I am grateful to the Center for Advanced Studies at Ludwig Maximilian University of Munich, Longview, and Effective Ventures for supporting this research.

References

Aizerman, Mark, and Andrew Malishevski. 1981. "General Theory of Best Variants Choice: Some Aspects." *IEEE Transactions on Automatic Control* 26 (5):1030–40. https://doi: 10.1109/TAC.1981.1102777.

Aizerman, Mark A. 1985. "New Problems in the General Choice Theory." *Social Choice and Welfare* 2 (4):235–82. https://doi.org/10.1007/BF00292690.

Al-Najjar, Nabil I., and Jonathan Weinstein. 2009. "The Ambiguity Aversion Literature: A Critical Assessment." *Economics & Philosophy* 25 (3):249–84. https://doi.org/10.1017/S026626710999023X.

Bartha, Paul, and C. Tyler DesRoches. 2017. "The Relatively Infinite Value of the Environment." *Australasian Journal of Philosophy* 95 (2):328–53. https://doi.org/10.1080/00048402.2016.1182196.

Bartha, Paul, and C. Tyler DesRoches. 2021. "Modeling the Precautionary Principle with Lexical Utilities." *Synthese* 199 (3):8701–40. https://doi.org/10.1007/s11229-021-03179-4.

Bengio, Yoshua, Stuart Russell, Elon Musk, Steve Wozniak, Yuval Noah Harari, Emad Mostaque, Andrew Yang, and et al. 2023. "Pause Giant AI Experiments: An Open Letter." https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

Berger, James O. 1985. *Statistical Decision Theory and Bayesian Analysis. 2nd ed*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4757-4286-2.

Boyer-Kassem, Thomas. 2017 a. "Is the Precautionary Principle Really Incoherent?" *Risk Analysis* 37 (11):2026–34. https://doi.org/10.1111/risa.12774.

Boyer-Kassem, Thomas. 2017 b. "The Precautionary Principle Has Not Been Shown to Be Incoherent: A Reply to Peterson." *Risk Analysis* 37 (11):2039–40. https://doi.org/10.1111/risa.12812.

Buchak, Lara. 2023. "Philosophical Foundations for Worst-Case Arguments." *Politics, Philosophy & Economics* 22 (3):215–42. https://doi.org/10.1177/1470594X231158662.

[18] $E + \Gamma$ might be thought to capture the relevant sorts of trade-offs at the first tier with E-admissibility.

[19] See Buchak (2023) for an argument that risk avoidance—"the idea that we should pay more attention to worse scenarios, even when we can assign sharp probabilities"—rather than ambiguity aversion provides a proper foundation for precautionary reasoning.

Caplan, Bryan. 2021. "An Ageless Hypothetical." https://www.econlib.org/an-ageless-hypothetical/.
Chambers, Christopher P., and M. Bumin Yenmez. 2017. "Choice and Matching." *American Economic Journal: Microeconomics* 9 (3):126–47. https://doi.org/10.1257/mic.20150236.
Danilov, Vladimir, and G. Koshevoy. 2005. "Mathematics of Plott Choice Functions." *Mathematical Social Sciences* 49 (3):245–72. https://doi.org/10.1016/j.mathsocsci.2004.09.001.
Deaton, Angus. 2013. *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton, NJ: Princeton University Press.
Dewey, John, and James Hayden Tufts. 1932. *Ethics. Revised ed.* New York: H. Holt and Company.
Fishburn, Peter C. 1973. *The Theory of Social Choice. Vol. 264*. Princeton, NJ: Princeton University Press.
Gärdenfors, Peter, and Nils-Eric Sahlin. 1982. "Unreliable Probabilities, Risk Taking, and Decision Making." *Synthese* 53 (3):361–86. https://doi.org/10.1007/BF00486156.
Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with Non-Unique Prior." *Journal of Mathematical Economics* 18 (2):141–53. https://doi.org/10.1016/0304-4068(89)90018-9.
Hansson, Sven Ove. 1997. "The Limits of Precaution." *Foundations of Science* 2 (2):293–306. https://doi.org/10.1023/A:1009671722406.
Harrison-Trainor, Matthew, Wesley H. Holliday, and Thomas F. Icard. 2016. "A Note on Cancellation Axioms for Comparative Probability." *Theory and Decision* 80 (1):159–66. https://doi.org/10.1007/s11238-015-9491-2.
Hill, Brian. 2020. "Dynamic Consistency and Ambiguity: A Reappraisal." *Games and Economic Behavior* 120 (C):289–310. https://doi.org/10.1016/j.geb.2019.12.012.
Kamran, Abbasi. 2020. "Covid-19: The Fatal Attraction of Herd Immunity." *BMJ* 370:m3714. https://doi.org/10.1136/bmj.m3714.
Lenman, James. 2000. "Consequentialism and Cluelessness." *Philosophy & Public Affairs* 29 (4):342–70. https://doi.org/10.1111/j.1088-4963.2000.00342.x.
Levi, Isaac. 1980. *The Enterprise of Knowledge*. Cambridge, MA: MIT Press.
Levi, Isaac. 1986. "The Paradoxes of Allais and Ellsberg." *Economics and Philosophy* 2 (1):23–53. https://doi.org/10.1017/S026626710000078X.
Levi, Isaac. 2008. "Convexity and Separability in Representing Consensus." In *Arguments for a Better World: Essays in Honor of Amartya Sen*, edited by Kaushik Basu and Ravi Kanbur, vol. 1, Ethics, Welfare, and Measurement, 193–212. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199239115.003.0012.
Luce, Robert Duncan, and Howard Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey*. Mineola, NY: Courier Dover Publications.
Lundgren, Björn, and H. Orri Stefánsson. 2020. "Against the De Minimis Principle." *Risk Analysis* 40 (5):908–14. https://doi.org/10.1111/risa.13445.
Mogensen, Andreas L. 2021. "Maximal Cluelessness." *Philosophical Quarterly* 71 (1):141–62. https://doi.org/10.1093/pq/pqaa021.
Mogensen, Andreas L., and David Thorstad. 2022. "Tough Enough? Robust Satisficing as a Decision Norm for Long-Term Policy Analysis." *Synthese* 200 (1):36. https://doi.org/10.1007/s11229-022-03566-5.
Moulin, Hervé. 1985. "Choice Functions over a Finite Set: A Summary." *Social Choice and Welfare* 2 (2):147–60. https://doi.org/10.1007/BF00437315.
Pedersen, Arthur Paul. 2009. *Pseudo-Rationalizability over Infinite Choice Spaces*. Tech. rep., Carnegie Mellon University.
Peterson, Martin. 2002. "What Is a De Minimis Risk?" *Risk Management* 4 (2):47–55. https://doi.org/10.1057/palgrave.rm.8240118.
Peterson, Martin. 2006. "The Precautionary Principle Is Incoherent." *Risk Analysis* 26 (3):595–601. https://doi.org/10.1111/j.1539-6924.2006.00781.x.
Peterson, Martin. 2017. "Yes, the Precautionary Principle Is Incoherent." *Risk Analysis* 37 (11):2035–38. https://doi.org/10.1111/risa.12783.
Plott, Charles R. 1973. "Path Independence, Rationality, and Social Choice." *Econometrica: Journal of the Econometric Society* 41 (6):1075–91. https://doi.org/10.2307/1914037.
Resnik, David B. 2021. *Precautionary Reasoning in Environmental and Public Health Policy. Vol. 86 of the International Library of Bioethics.* Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-70791-0.
Rott, Hans. 2001. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning. Vol. 42.* New York: Oxford University Press.

Sahlin, Nils-Eric. 2006. "Levi on Risk." In *Knowledge and Inquiry: Essays on the Pragmatism of Isaac Levi*, edited by Erik J. Olsson, Cambridge Studies in Probability, Induction, and Decision Theory, 87–96. New York: Cambridge University Press.

Seidenfeld, Teddy. 2004. "A Contrast between Two Decision Rules for Use with (Convex) Sets of Probabilities: -Maximin versus E-Admissibility." *Synthese* 140 (1–2):69–88. https://doi.org/10.1023/B:SYNT.0000029942.11359.8d.

Seidenfeld, Teddy, Mark J. Schervish, and Joseph B. Kadane. 2012. "Forecasting with Imprecise Probabilities." *International Journal of Approximate Reasoning* 53 (8):1248–61. https://doi.org/10.1016/j.ijar.2012.06.018.

Sen, Amartya. 1970. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78 (1):152–57. https://doi.org/10.1086/259614.

Sen, Amartya. 2004. *Rationality and Freedom*. Cambridge, MA: Harvard University Press.

Siniscalchi, Marciano. 2009. "Two Out of Three Ain't Bad: A Comment on 'The Ambiguity Aversion Literature: A Critical Assessment'." *Economics & Philosophy* 25 (3):335–56. https://doi.org/10.1017/S0266267109990277.

Sprenger, Jan. 2012. "Environmental Risk Analysis: Robustness Is Essential for Precaution." *Philosophy of Science* 79 (5):881–92. https://doi.org/10.1086/667873.

Steel, Daniel. 2015. *Philosophy and the Precautionary Principle*. Cambridge: Cambridge University Press.

Steel, Daniel, and Paul Bartha. 2023. "Trade-Offs and the Precautionary Principle: A Lexicographic Utility Approach." *Risk Analysis* 2023 (43):260–68. https://doi.org/10.1111/risa.13892.

Steele, Katie. 2006. "The Precautionary Principle: A New Approach to Public Decision-Making?" *Law, Probability and Risk* 5 (1):19–31. https://doi.org/10.1093/lpr/mgl010.

Stefánsson, H. Orri. 2019. "On the Limits of the Precautionary Principle." *Risk Analysis* 39 (6):1204–22. https://doi.org/10.1111/risa.13265.

Stewart, Rush T. Forthcoming. "Path Independence and a Persistent Paradox of Population Ethics." *Journal of Philosophy*.

Troffaes, Matthias C. M. 2007. "Decision Making under Uncertainty Using Imprecise Probabilities." *International Journal of Approximate Reasoning* 45 (1):17–29. https://doi.org/10.1016/j.ijar.2006.06.001.

United Nations. 1992. "Rio Declaration on Environment and Development." https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf. URL https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf.

Wingspread. 1998. "Wingspread Conference on the Precautionary Principle." https://www.sehn.org/sehn/wingspread-conference-on-the-precautionary-principle. URL https://www.sehn.org/sehn/wingspread-conference-on-the-precautionary-principle.

Appendix

Before proving the theorems, I will provide more formal and concise statements of the assumptions involved. For any alternative $X \in \mathcal{A}$ and event $E \subseteq S$, let $X[E] = \{x \in O : X(s) = x \text{ for some } s \in E\} \subseteq O$ be the image of E under X, the set of outcomes X has in states in E. For any $x \in O$, let $X^{-1}[x] = \{s \in S : X(s) = x\} \in \Sigma$ be the inverse image of $\{x\}$ under X, the set of all states for which alternative X has outcome x. Let $E_X^p = \left\{s \in S : \mathfrak{c}_p \in C\left(\left\{\mathfrak{c}_{X(s)}, \mathfrak{c}_p\right\}\right)\right\}$ be the set of states for which option X has an outcome that is not more choiceworthy than the catastrophic outcome p.

For theorem 1, first, we have $\mathrm{PP}(\alpha)_c$.

$$\text{Let } X \in \mathcal{A} \text{ be such that, for at least one } x_i \in X[S], \mathfrak{c}_p \in C\left(\left\{\mathfrak{c}_{x_i}, \mathfrak{c}_p\right\}\right). \text{ Then if } E_X^p \dot{\succ} E_Y^p, C(\{X, Y\}) = \{Y\}. \text{ If } E_X^p \dot{\sim} E_Y^p, \text{ then } C(\{X, Y\}) = \{X, Y\}. \tag{PP(α)}$$

Second, we have Cov_c. Recall that Cov_c says roughly that increasing the likelihood of a more choiceworthy outcome at the expense of a less choiceworthy outcome makes an act more choiceworthy overall.

> Let $X \in \mathcal{A}$ be such that $x_j \in X[S], C(\{\mathfrak{c}_{x_i}, \mathfrak{c}_{x_j}\}) = \{\mathfrak{c}_{x_j}\}$, and $X^{-1}[x_i] \dot{\succ} X^{-1}[x_j]$. Let X' be the alternative such that for any $s \in S$,
>
> $$X'(s) = \begin{cases} x_i, & \text{if } X(s) = x_j; \\ x_j, & \text{if } X(s) = x_i; \\ X(s), & \text{otherwise.} \end{cases} \qquad (\text{Cov}_c)$$
>
> Then, $C(\{X, X'\}) = \{X'\}$.

Now I present the (simplified) argument that these two assumptions are jointly inconsistent.

Proof of theorem 1

Proof. Let $X[S] = Y[S] = \{a, p, q\}$. Suppose that $X^{-1}[a] = Y^{-1}[a]$, $X^{-1}[p] = Y^{-1}[q]$, and $X^{-1}[q] = Y^{-1}[p]$, but $Y^{-1}[q] \dot{\succ} Y^{-1}[p]$. Because $E_X^p \dot{\sim} E_Y^p$, by $\text{PP}(\alpha)_c$, it follows that

$$C(\{X, Y\}) = \{X, Y\}. \tag{2}$$

Define Y' by

$$X'(s) = \begin{cases} q, & \text{if } Y(s) = p; \\ p, & \text{if } Y(s) = q; \\ Y(s), & \text{otherwise.} \end{cases}$$

Because $Y^{-1}[q] \dot{\succ} Y^{-1}[p]$, by (Cov_c),

$$C(\{Y, Y'\}) = \{Y'\}. \tag{3}$$

But $X = Y'$. Because C is a function, equations (2) and (3) are inconsistent. □

Now for theorem 2. Let E_* be an event that is not "negligibly unlikely." For any events $E, F \in \Sigma$, if E is *sufficiently more likely than* F, then there exists some event E_F^* such that, $E_F^* \dot{\succ} F$ and $E \dot{\succsim} E_F^*$. (The event E_F^* depends on F and may fail to exist.) First, we have $\text{PP}(\alpha)_c$.

> Let $X, Y \in \mathcal{A}$ be options such that there is exactly one $\hat{x} \in X[S]$ and exactly one $\hat{y} \in Y[S]$ such that $\mathfrak{c}_{\hat{x}}, \mathfrak{c}_{\hat{y}} \in \{p, q, \ldots\}$, and $\mathfrak{c}_{\hat{x}}, \mathfrak{c}_{\hat{y}} \in C(\{\mathfrak{c}_{\hat{x}}, \mathfrak{c}_{\hat{y}}\})$. Let $X^{-1}[\hat{x}], Y^{-1}[\hat{y}] \dot{\succsim} E_*$, and $Y^{-1}[\hat{y}] \dot{\succsim} E^*_{X^{-1}[\hat{x}]}$. Then, $C(\{X, Y\}) = \{X\}$ $\qquad (\text{PP}(\alpha)_c)$

Second, we have Arch_c.

> There are at least two alternatives $X, Y \in \mathcal{A}$ such that $X[S] = Y[S], X^{-1}[x], Y^{-1}[x] \dot{\gtrsim} E_*$ for all $x \in X[S] = Y[S]$, there is exactly one $\widehat{x}$ such that $\mathfrak{c}_{\hat{x}} \in \{p, q, \ldots\}, Y^{-1}[\hat{x}] \dot{\gtrsim} E^*_{X^{-1}[\hat{x}]}$, but $Y \in C(\{X, Y\}$. (Arch_c)

Recall that Arch_c does not specify how it is that $Y \in C(\{X, Y\})$, whereas Peterson assumes Y compensates for its greater likelihood to result in $\hat{x}$ by a greater likelihood to result in some "nice," noncatastrophic outcome.

Proof of theorem 2

Proof. By Arch_c, there exist at least two alternatives $X, Y \in \mathcal{A}$ such that $X[S] = Y[S]$, $X^{-1}[x], Y^{-1}[x] \dot{\gtrsim} E_*$ for all $x \in X[S] = Y[S]$, there is exactly one $\hat{x}$ such that $\mathfrak{c}_{\hat{x}} \in \{p, q, \ldots\}$, $Y^{-1}[\hat{x}] \dot{\gtrsim} E^*_{X^{-1}[\hat{x}]}$, but

$$Y \in C(\{X, Y\}). \tag{4}$$

However, because $Y^{-1}[\hat{x}] \dot{\gtrsim} E^*_{X^{-1}[\hat{x}]}$, by $\text{PP}(\alpha)_c$,

$$C(\{X, Y\}) = \{X\}. \tag{5}$$

Because $Y \neq X$, clearly, equations (4) and (5) are inconsistent.

Cite this article: Stewart, Rush T. 2025. "Deep Uncertainty and Incommensurability: General Cautions about Precaution." *Philosophy of Science*. https://doi.org/10.1017/psa.2024.40