

# On the Limits of the Precautionary Principle

H. Orri Stefánsson<sup>\*†</sup>

December 22, 2018

## Abstract

The Precautionary Principle (PP) is an influential principle of risk management. It has been widely introduced into environmental legislation, and it plays an important role in most international environmental agreements. Yet, there is little consensus on precisely how to understand and formulate the principle. In this paper I prove some impossibility results for two plausible formulations of the PP as a decision-rule. These results illustrate the difficulty in making the PP consistent with the acceptance of any trade-offs between catastrophic risks and more ordinary goods. How one interprets these results will however depend on one's views and commitments. For instance, those who are convinced that the conditions in the impossibility results are requirements of rationality may see these results as undermining the rationality of the PP. But others may simply take these results to identify a set of purported rationality conditions that defenders of the PP should not accept, or to illustrate types of situations in which the principle should not be applied.

KEY WORDS: Precautionary Principle; Environmental Policy; Catastrophic Risk

---

<sup>\*</sup>University of Copenhagen, Copenhagen, Denmark; Institute for Futures Studies, Stockholm, Sweden.

<sup>†</sup>Address: Faculty of Humanities, University of Copenhagen, Karen Blixens Plads 8, DK-2300 Copenhagen, Denmark; tel: 0046764281782; e-mail: orri.stefansson@hum.ku.dk.

# 1. INTRODUCTION: THE PRECAUTIONARY PRINCIPLE

The Precautionary Principle (PP) is a much discussed principle of risk management. The principle has been particularly influential on environmental policy—and has become part of e.g. the European Union’s environmental legislation and major international environmental agreements—but it has also played a prominent role in health policy and in the regulation of new technologies. A canonical and often cited version of the principle can be found in the United Nation’s (1992) *Rio Declaration*, whose Principle 15 states that:<sup>1</sup>

Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.

While the inclusion of the Precautionary Principle in environmental legislation and international agreements may have been a great political victory for environmentalists, it is hard to resist the thought that from a theoretical point of view, the principle is nothing but a truism, and is sufficiently vague to be open to interpretations that fit all plausible (but often conflicting) decision-rules.<sup>2</sup> Although “full scientific certainty” is open to different interpretations, it seems that no reasonable decision-rule would require such (“full”) certainty before recommending *cost-effective* measures in response to threats of “serious or irreversible damage”.

The Precautionary Principle has however been reformulated and sharpened since the publication of the Rio Declaration. For instance, Terje Aven (2010: 215) has more recently suggested that the following formulation captures how the principle is typically understood:<sup>3</sup>

---

<sup>1</sup>See <http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm> (22/05/2018).

<sup>2</sup>A decision-rule, as I am using the term, simply singles out an alternatives that no alternative is preferred to, if there is such an alternatives; otherwise it chooses nothing. So, for instance, if vanilla ice cream is preferred to all other items on an ice cream menu, then a decision-rule, based on that preference, singles out vanilla ice cream and nothing else. “Decision-rule” and other key terms are defined more carefully in the Appendix.

<sup>3</sup>Aven (2011a) has moreover usefully explicated the concept of *scientific uncertainty*; an explication that

The precautionary principle is the ethical principle saying that if the consequences of an activity could be serious and are subject to scientific uncertainties, then precautionary measures should be taken or the activity should not be carried out at all.

Aven's formulation of the Precautionary Principle is arguably stronger than that found in the Rio Declaration, as it *categorically* states that (what we might call) "cautionary" choices should be made—either precaution should be taken or the activity should not be carried out—when the conditions stated in the principle hold. In other words, his formulation does not limit the required cautionary choices to cost-effective ones.<sup>4</sup> Hence, Aven can be understood as being in agreement with those who interpret the principle as a rule that in certain circumstances prescribes cautious choices that can (and often do) conflict with the recommendations of standard cost-benefit analysis and decision theory (e.g. Gardiner, 2006, Steel, 2014).<sup>5,6</sup>

The aim of this paper is to present some problems for those who favor the Precautionary Principle understood along the above lines; that is, problems for those who favor the principle as a rule for choosing between risky (public or social) policies. In particular, I will be discussing problems for those who defend the PP as a partial decision-rule—or, equivalently, as a property of a decision-rule; i.e., a rule that only applies to certain situations (in particular, situations of scientific uncertainty) and prescribes cautious policy choices in those situations.

---

has been critically discussed by Cox (2011) and Vlek (2011), and which has been modified by Steel (2014), partly in response to worries like those raised by Cox. For further discussion of Aven's (2011a) proposal, see North (2011) and Aven (2011b). Moreover, Steel (2014) has, among others, recently extensively discussed the types of damage that the PP is meant to caution against. For a useful general recipe for sharpening the PP, see Sandin (1999) and Sandin, Peterson, Hansson, Rudén, & Juthe (2002: 290).

<sup>4</sup>On the other hand, it could be argued that the potential harm required for applying the PP is more severe on Aven's formulation than according to the Rio Declaration, if we assume that some "irreversible damage" is not serious.

<sup>5</sup>A referee for *Risk Analysis* questions this choice-rule-interpretation of Aven's formulation. So, to clarify, to prescribe some measure or (in)activity in certain situations is, as I understand it, to prescribe that certain choices are made in those situations.

<sup>6</sup>This is not to say however that Aven takes the PP to be inconsistent with risk analysis. In fact, he takes the latter to provide an input into the judgment about whether the conditions for applying the PP hold (Aven, 2010: 220; see also Aven, 2011a).

The problems in question stem from the fact that to avoid the charge of absolutism, most defenders of the Precautionary Principle, as a decision-rule, (explicitly) accept some trade-offs between catastrophic risks<sup>7</sup> and chances for more ordinary goods (see e.g. Steel, 2014). To make the PP consistent with such trade-offs, it would seem that it either has to include a proviso stating that some differences in catastrophic risks are small enough to overlook, or to include a proviso stating some threshold below which catastrophic risk is acceptable. For if it contains neither of these provisos, then any increase in catastrophic risk—no matter how small the increase and how insignificant the resulting risk—could outweigh any positive benefit associated with the increase.<sup>8</sup> But, as I will prove (with the help of technical conditions that I shall make explicit below), a PP that includes the first type of proviso violates one condition that (I contend) any good rule for selecting between risky policies should satisfy, while a PP that includes the second type of proviso violates another such adequacy condition.

Before we proceed, it might be worth emphasizing the scope (and limits) of the present article. First, some interpretations of the Precautionary Principle are unaffected by my argument. My aim is to logically analyze only one (quite widespread) interpretation of the PP, namely, an interpretation according to which the principle is to be understood as a rule for choosing between risky public (or social) policies. Various other interpretations of the PP have been proposed. Some for instance argue that the PP is meant to encourage precaution when setting up decision-problems (Steele, 2006, Peterson, 2017a: 5.2). Others take the PP to be a “meta-rule” for selecting decision-rules (for a discussion of this approach, see Steel, 2014: 2.3). Finally, some interpret the PP as an *epistemic* constraint on what to believe (see e.g. Peterson, 2007, 2017a: 5.3). These and other alternative interpretations will be set aside from now on.

---

<sup>7</sup>I will occasionally use the (informal) term “catastrophic risk” to refer to the type of danger that the PP encourages us to take measures against. However, as will become clear, such “risk” should not be interpreted as a (quantitative) probability.

<sup>8</sup>It is of course possible that provisos other than these two could succeed in making the PP non-absolutist. However, I am skeptical that such provisos would still be faithful to the thought behind (or the essence of) the principle, rather than turning it into something that is (according to proponents of the PP) too close to standard expected utility maximizing reasoning.

Second, as will become apparent, I will focus on rather simple versions, or special cases, of the Precautionary Principle as a decision-rule. However, I will argue (section 2.3) that what I say about these simple versions, or special cases, has implications for any fully general decision-rule-version of the principle.

Third, my argument has no implication for some *applications* of the Precautionary Principle, even if the principle is understood as a decision-rule. In particular, I shall be considering applications of the principle, as a decision-rule, in situations where one can at least compare some possible outcomes in terms of their relative likelihood, without knowing their precise probabilities. I will follow the convention of calling such predicaments *situations of uncertainty*, but I shall argue (section 2.2) that they are perfectly consistent with what is often called “scientific uncertainty” in the literature on the Precautionary Principle. As we shall see, some, but not all, defenders of the principle want to apply it, as a decision-rule, in situations of uncertainty. The results I will present illustrate certain problems that one is faced with when trying to apply the principle in such situations. But these results do not have any implications for applications of the principle in what is often called situations of “complete ignorance”, i.e., in situations where one is not able or willing to make any comparative likelihood judgments.<sup>9</sup>

The above observations illustrate that one can interpret the results of this paper as delineating the role and scope of the Precautionary Principle: *if* one accepts that the principles in the impossibility results that I prove are adequacy conditions on rules for choosing between e.g. public policies, then my results show that one should not interpret the PP as a rule for choosing between policies in situations of uncertainty. In contrast, those who interpret the PP as a decision-rule that can be applied in situations of uncertainty (e.g. Steel, 2014), can interpret the results of this paper as illustrating what other principles they must reject. In particular, such scholars could interpret the results of this papers as clarifying which principles of standard cost-benefit analysis and

---

<sup>9</sup>It might be worth mentioning that I subscribe to the view that we are almost never in situations of complete ignorance. But I will not argue for that view here.

decision theory they must reject (for the situations in which they want the PP to hold).<sup>10</sup> Finally, one could resist the importance of these results by questioning the technical assumptions or simplifications used in proving them.

The rest of this paper is structured as follows. I will start by discussing (in section 2) a formulation of the Precautionary Principle that includes a proviso implying that some differences in catastrophic risks are small enough to overlook, and which is essentially a formulation that Martin Peterson (2006) has proposed and Steel (2014) has endorsed. The first impossibility theorem I prove strengthens a theorem that Peterson (ibid.) proves in two respects: first, it dispenses with one of the conditions of Peterson's theorem, second, it weakens another of Peterson's conditions. The latter weakening means that my argument avoids some recent objections to Peterson's argument (Steel *ibid.*, Boyer-Kassem, 2017a). The new theorem states that if we assume a weak (Archimedean) principle about trade-offs between catastrophic risk and more ordinary goods, then the version of the PP that includes the difference proviso violates (given an additional technical condition) the *Transitivity* principle, which states that if  $X$  is preferred to  $Y$  and  $Y$  is preferred to  $Z$  then  $X$  is preferred to  $Z$ .<sup>11</sup>

My second formulation of the Precautionary Principle (discussed in section 3) includes a proviso implying that catastrophic risk can be ignored if it falls below some (greater than negligible) threshold. While this version of the PP does not violate *Transitivity*, it does (again, given the technical condition plus a weak Archimedean principle) violate another condition that I call *Sub-alternative Dominance* (SAD). Intuitively, this condition states that when comparing alternatives along different dimensions, one's ranking of alternatives within the different dimensions should be reflected by one's overall ranking of the alternatives.

Violating *Sub-alternative Dominance* might be something that supporters of the

---

<sup>10</sup>Daniel Steel has confirmed, in personal communication, that the latter is indeed how he interprets the results of this paper. A referee for *Risk Analysis* suggested this interpretation too, which is in line with how Boyer-Kassem (2017b) interprets Peterson's (2006) impossibility results for the PP.

<sup>11</sup>In what follows, the preferences we will be discussing should be understood as being those of a "social planner" (i.e., a policy-maker of some sort).

Precautionary Principle could live with, for instance since it could be doubted that SAD is a coherence constraint (like Transitivity). But, as I explain in section 3, violating SAD comes with practical costs in the present context: it means that the “social planner” cannot conduct “local” risk evaluations (e.g. for different policies, products, or regions), but has to conduct one “global” risk evaluation where she evaluates simultaneously all policies or products that she will ever consider. Obviously, that is not a feasible task.

## **2. IMPOSSIBILITY RESULT FOR DIFFERENCE PROVISIO**

### **2.1. First formulation**

Although there is still considerable scope for disagreement when it comes to precisely how to understand the Precautionary Principle, interpreted as a decision-rule, a logical analysis of the principle can take as starting point some (weak or uncontroversial) implication of it, that those who defend the principle as a decision-rule could agree on. That is the strategy favored by Martin Peterson (2006), who suggests the following principle (a formal statement of which is given in the Appendix) to that end:

**PP( $\delta$ ):** 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.

The terms “fatal”, “negligibly unlikely” and “sufficiently more likely” are all left (intentionally) vague. But the thought is that the Precautionary Principle is meant to caution against outcomes whose badness or undesirability is above some (vague, imprecise) threshold; often called “catastrophic” outcomes, but in Peterson’s framework called “fatal” outcomes (I will use these two terms interchangeably). Moreover, the idea is that for the principle not to be too extreme (or absolutist), it should only prescribe

caution when the likelihoods (explained below) of the possible fatal outcomes are non-negligible, and it should not require us to take precautionary measures (i.e., to prefer the latter act in  $PP(\delta)$ ) unless such measures reduce the likelihood of the fatal outcome by some “sufficient” (vague, imprecise) amount.

I will resist the temptation to speculate about whether the signatories of the Rio Declaration would accept  $PP(\delta)$ . Instead, I will simply note that Steel (2014: 40), who has recently defended the PP as a decision-rule, admits that  $PP(\delta)$  is a consequence of his interpretation of the PP. Steel however responds to Peterson’s impossibility result by rejecting one of its premises. But as we shall soon see (section 2.4), Steel’s objection does not get the PP off the hook.

## 2.2. Likelihoods

It is important to note that the aforementioned likelihoods should not be interpreted as quantitative probabilities. As previously mentioned, the Precautionary Principle, on the interpretation that I will be assuming (following, for instance, Peterson, 2006), is meant to apply to what I call situations of uncertainty. As the reader may recall, these are situations where a decision-maker can rank at least some of the possible outcomes of the alternatives (or acts; I will use these terms interchangeably) she is considering in terms of their likelihood—i.e., in terms of how confident she is that they will materialize if the alternative is chosen—without necessarily knowing their probabilities.<sup>12</sup>

As an illustration of the types of likelihood judgments that I have in mind, imagine that a decision-maker is considering using solar geoengineering to reduce global warming. Although the decision-maker (or the analysts she consults) might not know precisely by how much that will reduce the probability of a global warming catastrophe, she might have some reliable evidence and/or general theories which suggest that these measures will somewhat reduce the likelihood of such a catastrophe. Hence, the likelihood judgment (although “subjective”) is what epistemologists would typically

---

<sup>12</sup>Note however that for instance Steel (2014: ch. 5) takes the principle to hold even in some situations where quantitative probabilities are known, which of course entails that the likelihoods are known too.

call *justified* (see e.g. Steup, 2018). At the same time, she might judge that since solar geoengineering is such a radical and novel intervention, it may increase the likelihood of some *yet unforeseen* undesirable outcome, without being able to quantify that increase. However, even if she is unable to make the relevant differences in the likelihoods precise, she might judge that the increased likelihood of an unforeseen undesirable outcome caused by solar geoengineering is more than offset by the decrease in the likelihood of a global warming catastrophe.

The assumption that the decision-maker of interest (or the analysts she consults) forms likelihood judgments of the above kind raises two related questions. First: Is this assumption consistent with assuming that the decision is being made in what is in the literature on the Precautionary Principle called a situation of *scientific uncertainty* (which is considered a pre-condition for the applicability of the principle)? Second: If the decision is made in a situation of scientific uncertainty, is it then really justifiable to base the decision on such likelihood judgments? After all, in such situations, it might seem that these judgments “should not be given much weight, as the knowledge supporting them is so poor”, as a referee puts it. Let us consider these questions in turn.

For concreteness, I shall focus on Terje Aven’s explication of scientific uncertainty, but what I say holds for other explications too (for instance, the one defended by Steel, 2014). Roughly, Aven suggests that there is scientific uncertainty when it comes to the effects of, say, solar geoengineering, if either no accurate prediction model of the effects of solar geoengineering can be established, or such a model can be established but there is sufficiently great uncertainty about the model’s input parameters (Aven, 2011a).

Now let’s compare Aven’s explication of scientific uncertainty to our imagined decision-problem three paragraphs back. As I understand it, there is, in this imagined case, scientific uncertainty of the less severe kind (i.e., what Aven *ibid.* calls “category II\*”) when it comes to the effects of solar geoengineering on climate change. That is, there is an accurate predication model but there is great uncertainty about some of its input parameters. Or at the very least, the description of the problem is consistent with

such uncertainty, so we can simply stipulate for the sake of the argument that such uncertainty holds in that case. However, there is the more severe type of scientific uncertainty (“category III\*”) when it comes to the overall effects of solar geoengineering, I take it, since we cannot even specify a full space of possible outcomes from solar geoengineering (nor the full state space that might give rise to these outcomes); and hence, we have no accurate prediction model of the overall effects of solar geoengineering.

But even though the decision in question is by assumption made in a situation of scientific uncertainty, it does not follow that there is no basis on which the decision-maker (or the analysts) can form likelihood judgments. For subjective judgments of likelihoods need not, of course, be based on complete and accurate prediction models. They could instead be based on less complete background knowledge and expert judgment, that for instance includes some data, incomplete models and general theories (see e.g. Aven, 2013 and Steel, 2014).<sup>13</sup>

What about the second question: Are we justified in basing decisions on likelihood judgments in situations of scientific uncertainty? Arguably, we should not give such judgments *as much* weight as we give probabilities based on, say, long run frequencies or probabilities outputted by accurate prediction models. If we have such well-corroborated probabilities, for the outcomes of interest, then most people would agree that the decision should maximize expected value of some sort (e.g. expected utility, risk-weighted expected utility, etc.). In contrast, when the likelihood judgments are made in situations of scientific uncertainty—and, in particular, if the potential outcomes could be catastrophic—then we should not, if proponents of the Precautionary Principle are right, seek to maximize expected value of any sort, even if we can turn these judgments into sufficiently precise probabilities (Steel, 2014). One way to interpret this resistance to expected value maximization, is that we decide to not give full weight to these likelihood judgments.

---

<sup>13</sup>As Aven (2013) notes, however, it is crucial that the analyst report what their judgments (or their expressed probabilities) reflect; e.g. whether they are based on long run (or hypothetical) frequencies or on less complete information.

But the above does not mean that we should give *no* weight to likelihood judgments in situations of scientific uncertainty. After all, these judgments can, as previously discussed, be based on valuable (albeit incomplete) empirical evidence as well as expert judgment or general theories. So, it would seem imprudent—and unreasonable—to disregard such judgments altogether, for that would be to disregard information that we would expect to be useful in furthering our aims. Instead, we should arguably try to combine these judgments with some more cautious decision-rule (i.e., more cautious than the standard expected value rules), such as a decision-rule based on the Precautionary Principle—if a coherent decision-rule based on the PP can be formulated.

In what follows, I will be assuming that the likelihood judgments in question satisfy all the conditions for being what is called a *qualitative probability* (Krantz, Luce, Suppes, & Tversky, 1971: 204-205), except for completeness. A formal statement of these conditions can be found in the Appendix. Importantly, being a qualitative probability does not suffice for representability by a probability function nor even by a set of probability functions (ibid.: 205-206). So, while these likelihoods have a considerable structure, they need not have all the qualitative structure entailed by probabilities, and should not be assumed to be quantitative probabilities. Nevertheless, even setting aside completeness, peoples' judgments often fail to have the structure of qualitative probabilities (as for instance illustrated by the "Conjunction Fallacy"; see e.g. Kahneman & Tversky, 1982). Hence, contrary to what a referee suggests, the assumption that the decision-maker of interest can form relative likelihood judgments, thus understood, is far from being trivial.

### **2.3. Some objections**

Before explaining the trouble with  $PP(\delta)$ , I will address three potential objections to the claim that  $PP(\delta)$  is a good explication of the Precautionary Principle. (These objections could also be raised against the explication considered in the next section.) The three objections illustrate that neither the versions of the principle to be discussed nor the

underlying framework to be employed are sufficiently general to be applicable in all situations in which one might like to apply the PP. Nevertheless, if one accepts that  $PP(\delta)$ —or  $PP(\epsilon)$ , which I discuss in the next section—really is implied by the general principle, in the type of situation to which  $PP(\delta)/PP(\epsilon)$  is limited, then that is sufficient to generate trouble for the PP. After all, if an implication of a more general principle is false, then so is the general principle. For a truth cannot imply a falsehood.

Now a defender of the PP could of course respond that I still have not shown that, say, other special cases of the PP—cases that are limited to types of situations that I do *not* discuss—are problematic. Furthermore, they might argue, I have given them no reason for why they should not continue to use the PP in such situations; and, in fact, they might claim that they only intended the PP for such situations. While this is a perfectly valid objection, I will not try to address it, except by re-emphasising that my aim is just to explore the special cases of the PP that I describe below, and, by implication, to explore any more general principle that implies these special cases.<sup>14</sup>

Let's then consider the three objections. First, Thomas Boyer-Kassem (2017a, 2017b) has recently argued that  $PP(\delta)$  cannot be an acceptable explication of the Precautionary Principle since it “considers only the most likely fatal outcome for each action and disregards other fatal outcomes” (2017a: 2031). Boyer-Kassem's own purported counterexample to  $PP(\delta)$  is, I think, not so hard to deal with, since it involves different states of the world that, for each act, result in the same outcome. Hence, while he insists that the states are different, they should be merged, if we treat acts as functions from states to outcomes, as Peterson (2006: 600) does, and as is typically done by decision theorists (at least by those working in the dominant tradition that is often associated with Leonard Savage, 1954). For according to that treatment, states are, to put it informally, differentiated by their capability of producing different outcomes. So, two states that, for each act, deliver the same outcome, are treated as the same state. Nevertheless, as Boyer-Kassem (2017b) points out, one could construct examples involving a number of

---

<sup>14</sup>I thank a referee for pressing me on this issue.

different possible fatal outcomes, which might cause trouble for  $PP(\delta)$ , due to the fact that the principle only considers the most likely fatal outcome.

One possible response to the above worry would be to suggest that intuitively different outcomes that are all classified as “fatal” should be grouped together as one outcome, which contains uncertainty as to precisely how bad it is. A problem with that response is that  $PP(\delta)$  then fails to discriminate between alternatives that intuitively should be discriminated between, such as two alternatives that have the same likelihood of resulting in the fatal outcome(s) but where the fatal outcome(s) that may result from one of the alternatives is (on balance) worse than the fatal outcome(s) that may result from the other.

A better response is to admit that  $PP(\delta)$  is a very simple implication of the full principle (understood as a decision-rule). For instance, one can think of  $PP(\delta)$  as the special case where there is only one possible fatal outcome associated with each alternative under consideration (and the same holds for  $PP(\epsilon)$  which I discuss in the next section). In other words, we assume that the full principle implies  $PP(\delta)$  for such simple situations. And note that this may suffice to undermine the fully general principle: If this implication of the Precautionary Principle is false, then the principle itself is false too (even though other special cases of the principle may be true).

Now, it might not always be clear whether a decision-problem should be described as one involving only one fatal outcome or multiple fatal outcomes. And, as Boyer-Kassem (2017a) has shown, applying  $PP(\delta)$  in such cases might result in different choices depending on different ways of describing the problem. To put it more precisely,  $PP(\delta)$  will deliver different verdicts depending on different ways of partitioning the state-space. Boyer-Kassem takes this to show that  $PP(\delta)$  cannot be an adequate explication of the Precautionary Principle, since he seems to think that the choices recommended by the latter should not depend on how the decision-problem is formulated. However, even decision theories as idealized as Leonard Savage’s (1954) may recommend different courses of action depending on how the state-space is partitioned. In fact, among the

traditional theories, only the one developed by Richard Jeffrey (1965), and theories based on it, are what is known as *partition invariant* (for a discussion, see e.g. Joyce, 1999: 121-122). Hence, we should not expect that any decision-rule version of the PP will be partition invariant. In other words, we should expect that the choices it recommends may depend on different—and possibly even equally natural or “correct”—ways of partitioning the state-space.<sup>15</sup> Nevertheless, I take Boyer-Kassem’s argument to be a useful reminder that PP( $\delta$ ) should only be applied to situations where it is intuitively clear that there is only one fatal outcome at stake.

Second, a referee for *Risk Analysis* points out that the Precautionary Principle might be applied in situations where the scientific uncertainty means that “subjective likelihood statements are not really worth much”. (Boyer-Kassem, 2017a makes a similar point.) In contrast, as mentioned above, I will be assuming an interpretation of the principle according to which it is applied in situations where meaningful likelihood (i.e., comparative confidence) judgments can be made. Again, this may mean that PP( $\delta$ ) is a special case of the full PP; i.e., a case limited to situations where such comparative judgments can be made. Consequently, my discussion will have little if any bearing on precautionary choices in situations where such comparative judgments cannot be meaningfully made.

Nevertheless, since notable recent defenders of the PP, such as Steel (2014: ch. 5), take the principle to be sufficiently general to have implications for situations where comparative likelihood judgments can be made (and even for some situations where probabilities are known), I believe that my discussion is highly relevant to the general principle. After all, I argue that an implication of the general principle leads to trouble, which is of course trouble for the general principle too. Moreover, as Peterson (2017b) points out, impossibility results based on the framework employed in this paper hold even if the decision-maker does not *know* how to make the likelihood comparisons; the results only require that such comparisons *can* be made. Finally, as explained in

---

<sup>15</sup>Peterson (2017b) offers a similar reply, appealing to a formal result from Peterson (2009).

some detail above, the assumption that likelihood judgments can be made is perfectly consistent with the notion of “scientific uncertainty” which figures prominently in most statements of the Precautionary Principle.

Third, Boyer-Kassem (2017a: sec. 3) complains that Peterson’s (2006) framework, parts of which I take on board, assumes that the decision-maker is aware of all outcomes that the acts she is considering could result in. Strictly speaking, however, the impossibility results do not rest on the decision-maker being fully aware (no more than they rest on the decision-maker knowing the relevant likelihoods). Nevertheless, Boyer-Kassem is correct in pointing out that the framework in question does not model what is sometimes called situations of (conscious) *unawareness* (see e.g. Walker & Dietz, 2011, Grant & Quiggin, 2013, Karni & Vierø, 2013). The same holds for PP( $\delta$ ). This limits the scope of Peterson’s framework and the principles and results he discusses, as Boyer-Kassem correctly points out, and could seem unfortunate since we might want to apply the Precautionary Principle in situations of conscious unawareness. However, I think we should not take this to undermine the importance of the results in question, since no plausible and precise decision-rule—precautious or not—has yet been formulated for situations of conscious unawareness.

#### **2.4. Trade-offs and (im)precision**

The trouble with PP( $\delta$ ) partly arises due to the fact that those who endorse precautionary-type reasoning typically accept *some* trade-offs between increased likelihood of a fatal outcome and increased likelihood of more ordinary goods (see e.g. Steel, 2014: ch. 4). To take an example, there might be some increase in the risk of a catastrophic outcome that could be made up for by sufficiently increasing the chance of people enjoying higher welfare. To make this more concrete, consider the catastrophic (even existential) risks from artificial intelligence that many scholars have warned about. Even knowing about these risks, most people happily endorse both basic and applied AI research—which plausibly (nonnegligibly) increases the likelihood of the catastrophes in question—due

to the great benefits that AI promises, for instance by reducing traffic fatalities, pollution and congestion, and by improving the accuracy of medical diagnoses.

Peterson (2006: 599) suggests a principle logically equivalent to the following, to capture the thought behind trade-offs like those I have been alluding to:

**Archimedes (A):** Some (nonnegligible) increase in the relative likelihood of a fatal outcome can be counterbalanced precisely by the relative likelihood of a nonfatal outcome being increased in relation to a strictly worse nonfatal outcome.

While A may, on the face of it, seem like a weak and plausible principle, there are reasons to worry about the assumption that changes in the likelihood of fatal outcomes can be counterbalanced *precisely* by changes in the relative likelihood of nonfatal outcomes. The Precautionary Principle is typically meant to be used in situations where quantitative information about probability (and perhaps utility) is lacking, i.e., in situations where agents have no choice but to reason with qualitative information. In such situations, Steel (2014: 41-42) argues that Peterson's Archimedean condition is implausible. He asks the reader to imagine a situation where a decision-maker has only a coarse-grained ranking of outcomes in terms of their utility and "credibility" (i.e., what Peterson calls "relative likelihood"); on one hand a ranking of outcomes into the categories "excellent", "poor", "good", and "catastrophic"; on the other hand a ranking according to the categories "strong evidence", "some evidence", "minimal evidence", and "merely possible". Then he asks:

[W]hat 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. For to know how to answer it, one would need some quantitative measure of the increases and decreases in credibility and utility involved. Is the utility of the excellent outcome twice that of the good outcome or

three times? Similarly, how much worse is a catastrophic outcome than a poor one? And what is the exact degree of difference between “very strong evidence” and “strong evidence” and so on? In short, A is sensible only given quantitative measures of credibility and utility that are, by definition, absent in reasoning with qualitative [*sic*] information. To insist that a theory of decision with qualitative information be committed to A is to unreasonably demand that such theories draw arbitrary distinctions. (42)

Although Steel may have a point, his observation does not get the Precautionary Principle, as a decision-rule, off the hook. Even a defender of the PP should, I contend, accept that it is sometimes the case that an “increase of credibility of catastrophe relative to poor would [be *at least*] offset [by] the advantage accruing from the increase of the credibility of excellent relative to good”. In other words, we can, without quantitative information, be sure that there is (i) some increase in the credibility of catastrophe relative to poor and (ii) some increase in the credibility of excellent relative to good, such that if we compare, say, a policy  $P_1$  with another policy  $P_2$ , that only differs from  $P_1$  in terms of both increase (i) and increase (ii), then  $P_2$  is *at least* as good as  $P_1$ . And this would be enough to generate trouble for the PP. For as I prove in the Appendix, Peterson’s impossibility theorem can be strengthened by replacing his Archimedean condition with the following (strictly logically weaker) condition, which avoids Steel’s objection:

**Weak Archimedes (WA):** Some (nonnegligible) increase in the relative likelihood of a fatal outcome can be *at least* counterbalanced by the relative likelihood of a nonfatal outcome being increased in relation to a strictly worse nonfatal outcome

Here is a pair of examples that illustrates the difference between A and WA. According to the first principle, it is possible to (nonnegligibly) increase the likelihood of a catastrophic outcome but still leave things *precisely* as good, or desirable, as they are,

by sufficiently increasing the likelihood of people enjoying higher welfare than they otherwise would. According to the latter principle, it is possible to (nonnegligibly) increase the likelihood of a catastrophic outcome but still leave things *at least* as good, or desirable, as they are, by sufficiently increasing the likelihood of people enjoying higher welfare than they otherwise would.

## 2.5. Incommensurability

In addition to avoiding Steel's objection, the weakening of the Archimedean condition avoids Boyer-Kassem's (2017a) worry that Peterson's A assumes commensurability between fatal outcomes and nonfatal ones; an assumption that he takes to contradict the motivation behind the Precautionary Principle.

Boyer-Kassem is right in thinking that if fatal outcomes really are incommensurable in value to nonfatal outcomes, then Peterson's A is false. In general, if a good  $G_1$  is incommensurable in value to good  $G_2$  which is comparable to, and, say, better than  $G_3$ , then there typically are no probabilities,  $p, q, p', q'$  (where  $p \neq p', q \neq q'$ ) such that a gamble that confers a probability of  $p$  on  $G_1$ , a probability of  $q$  on  $G_2$ , and a probability of  $1-(p+q)$  on  $G_3$  is *precisely* as good as another gamble that confers a probability of  $p'$  on  $G_1$ , a probability of  $q'$  on  $G_2$ , and a probability of  $1-(p'+q')$  on  $G_3$ .

So, if Boyer-Kassem is right, then Peterson's A is false. But what about WA? On what I think is the most popular account of "incommensurability"—according to which incommensurability between two goods neither means complete incomparability nor that one good always dominates the other (see e.g. Chang, 2013)—WA can be true even if fatal and nonfatal outcomes are incommensurable in value.<sup>16</sup> In general, if two goods are incommensurable, then for *some* ranges of quantities<sup>17</sup> of each, there might be no

---

<sup>16</sup>To take another example, according to Steel's (2014: 113-114) definition of commensurability, WA does not entail that valuations of nonfatal and fatal outcomes are commensurable. He assumes that valuations of goods in set  $S$  are commensurable "just in case there is some unit of value  $v$  [...] such that for any  $s \in S$  there is some finite quantity  $q$  of  $v$  such that the person is indifferent between  $s$  and  $q$ ." In contrast, WA would at most entail that for any outcome  $x$ , where  $x$  could be either fatal or nonfatal, there is some quantity  $q$  of  $v$  that the person considers *either* to be at least as great as the value of  $x$  or no greater than the value of  $x$ .

<sup>17</sup>I am not assuming that we have a numerical representation of the utility of the good in question; the

fact of the matter as to whether one quantity is better than, worse than, or equally good as the other. But, I contend, the same need not hold for *any* quantity of each. However, even if for any quantity  $m > 0$  of good  $g_1$  and any quantity  $n > 0$  of good  $g_2$  it is false that  $m$  quantities of  $g_1$  are *precisely* as good as  $n$  quantities of  $g_2$  (which I think could be true of many goods), it may very well still be true that some quantities of  $g_1$  are *at least* as good as some quantities of  $g_2$  (see Chang, *op. cit.*).

To take an example, suppose that freedom and pleasure are incommensurable ingredients of a good life and imagine that we have found a way to measure, on an (ordinal) scale from 1 to 10 (which I will call “units”), a person’s overall freedom, and that we have similarly found a way to measure, on an (ordinal) scale from 1 to 10, how pleasurable someone’s life is as a whole. Then there might be no fact of the matter as to how to rank the following two lives in terms of how good they are: one that contains 4 units of pleasure and 6 units of freedom and another one that contains 6 units of pleasure and 4 units of freedom; moreover, it might not be true of any two (distinct) distributions of units of pleasure and freedom,  $d_i$  and  $d_j$ , that a life containing  $d_i$  is equally good as a life containing  $d_j$ ; but it might still be true that a life containing 10 units of pleasure and 2 units of freedom is at least as good as a life containing 2 units of pleasure and 3 units of freedom.<sup>18</sup>

It is worth noting, however, that Boyer-Kassem (2017a: 2030-2031) seems to have a stronger notion of incommensurability in mind than the one suggested above, namely, one according to which *no* increase in the likelihood of a fatal outcome can be compensated for by *any* change in the likelihood of a nonfatal outcome, i.e, an interpretation that simply negates what the Archimedean condition postulates (thanks to a referee for pressing me on this point). He claims support for this view by noting that the thought underlying the distinction between fatal outcomes and nonfatal ones is that a

---

quantities could e.g. be numbers of dollars or numbers of years in good health (thanks to Martin Peterson for encouraging me to clarify this).

<sup>18</sup>Dominance principles that entail that, say, a life containing 4 units of pleasure and 5 units of freedom is at least as good as a life containing 3 units of pleasure and 4 units of freedom, are not strong enough to sanction the kind of trade-off reasoning that the Archimedean conditions under discussion are meant to formalize.

fatal outcome is so bad that even a large—or “enormous”, to use Peterson’s (2006: 597) terms—benefit from nonfatal outcomes cannot make up for it.

Now “large” and “enormous” are vague terms and one could very well accept the aforementioned thought but still hold that *some* (maybe “gigantic”?) benefit from nonfatal outcomes can make up for *some* harm from fatal outcomes. Nevertheless, it is of course possible that some defenders of the Precautionary Principle will simply insist that no trade-offs between chances for fatal and non-fatal outcomes are ever acceptable, even if support for such insistence cannot be found in the aforementioned distinction. Moreover, as a referee points out, “someone who plainly says that environmental goods cannot be compared with money, full stop, would not agree with the proposed solution”, i.e., would not even accept WA. In response, I must say that the view that *no* amount of environmental good could be compared to *any* monetary amount, and the view that *no* trade-offs between chances for fatal and non-fatal outcomes are *ever* acceptable, both strike me as being simply too implausible to deserve serious discussion. At any rate, I admit that I have little hope of convincing those who hold such absolutist views. Fortunately, many defenders of the PP would reject such views too. In fact, as Steel points out, “defenses of the PP often dismiss absolutist interpretations as unfair distortions” (2014: 84).

But perhaps defenders of the Precautionary Principle can develop an argument against even the weak Archimedean principle that does not lead to the absolutist views discussed above. One reason why defenders of the PP might want to resist WA is that although the condition is, I contend, satisfied by any acceptable rule for choosing between risky public policies, some might find that the PP and Archimedean conditions in general “express conflicting views”, as Boyer-Kassem (2017a: 2039) puts it. If that is the case, then some defenders of the PP might be perfectly happy with giving up even WA. Consequently, if this is true, then the results of this paper will not be worrying for those defenders of the PP. Instead, these results could simply be interpreted as contributing to an understanding of, first, what a defender of the PP must commit

to, and, second, how precisely the PP differs from standard cost-benefit analysis and decision theory.

Another reason why defenders of the Precautionary Principle might, and arguably should, try to resist even WA, is that it turns out that one cannot simultaneously endorse the PP, as formulated above, and accept trade-offs of the type I have been discussing, as formulated by WA, while satisfying Transitivity. As you may recall, Transitivity states that, for any alternatives  $X$ ,  $Y$  and  $Z$ , if one (weakly) prefers  $X$  to  $Y$ , and (weakly) prefers  $Y$  to  $Z$ , then one (weakly) prefers  $X$  to  $Z$  (which means that one does *not* strictly prefer  $Z$  to  $X$ ).<sup>19</sup> I will get back to this impossibility result in subsection 2.7, after having discussed one more assumption of the result.

## 2.6. Increasing likelihoods

Previously I mentioned that a technical condition has to be assumed to prove the above mentioned impossibility result. It is time to make explicit this assumption, which is implicit in Peterson's (2006) argument too.<sup>20</sup> Recall that the weak Archimedean condition essentially ensures that the likelihoods are sufficiently fine-grained such that one can increase the likelihood of a catastrophe by such a small amount that it can be made up for by, say, some gigantic benefit in terms of non-catastrophic outcomes. The yet unstated technical condition can be understood as the flip-side of the Archimedean condition: It ensures that one can also (in principle) increase the likelihood of a catastrophic outcome by so much that it becomes "sufficiently more likely", as defined by  $PP(\delta)$ , than what is previously was, unless the outcome was already (almost) certain to occur. Hence, this technical condition ensures that it is (in principle) possible to increase the likelihood of a catastrophic outcome by so much that a defender of the Precautionary Principle

---

<sup>19</sup>In fact, something strictly stronger holds: A decision-rule which satisfies both WA and  $PP(\delta)$  violates *Acyclicity*, (one version of) which states that if one (weakly) prefers  $X$  to  $Y$ , and (weakly) prefers  $Y$  to  $Z$ , then one does not (strictly) prefer  $Z$  to  $Y$ . Transitivity entails Acyclicity, but Acyclicity only entails Transitivity if the preference relation is complete.

<sup>20</sup>Thanks to a referee for *Risk Analysis* for making me see the need to include an explicit discussion of this condition in the main text, rather than only discussing it in the Appendix as I did in a previous version of this paper.

will think that no benefit in terms of non-catastrophic outcomes could make up for the increase. (The assumption does not, however, imply that one can increase the likelihood of a catastrophe by any precise amount.)

To put it more accurately, the assumption is that if there is an alternative  $X$  that could result in some catastrophic outcome  $a$ , then as long as the catastrophe is not already (almost) certain to occur, there is some other *logically possible* (but perhaps infeasible) alternative  $X'$  that is sufficiently more likely than  $X$  to result in  $a$ . A natural interpretation of this assumption, given how it is formalized in the Appendix, is that there are more states of the world under which  $X'$  would result in  $a$  than there are states of the world under which  $X$  would result in  $a$ . For instance, if  $X$  is the use of some chemical  $c$  that might lead to natural catastrophe  $a$ , then  $X'$  might differ from  $X$  in (only) containing a greater amount of  $c$ . Hence, it is natural to assume that, say, the number of states in which some natural system might be such that use of  $c$  results in catastrophe  $a$  is greater if alternative  $X'$  is chosen than if alternative  $X$  is chosen.

This technical condition strikes me as being an innocent assumption. In particular, it seems to be something that those who push for the Precautionary Principle must accept. For if it were impossible to increase the likelihood of a catastrophe sufficiently to trigger the principle, then the PP would always be trivially satisfied. But the condition should be acceptable to detractors of the PP too, since it only says that whatever counts as a “sufficient” increase in likelihood, as defined by  $PP(\delta)$ , it is possible to increase the likelihood of a catastrophe by at least that much. Hence, I will from now on typically omit to explicitly mention this technical condition and simply take it for granted as a background assumption. However, the possibility of resisting the results of this paper by refuting this technical condition would certainly be worth examining.

## **2.7. The importance of the impossibility**

Back then to the importance of the aforementioned impossibility result. It is relatively uncontroversial that no good decision-rule can fail to be transitive (but see e.g. Temkin,

2012 for some arguments against the orthodoxy).<sup>21</sup> Therefore, the upshot of the aforementioned result is that (given the discussed technical condition) no good decision-rule can simultaneously satisfy  $PP(\delta)$  and WA—which could be taken as evidence that the Precautionary Principle and Archimedean conditions do indeed express conflicting views. Hence, if I am right that no good decision-rule violates WA, then it follows that no good decision-rule can satisfy  $PP(\delta)$ . So, if  $PP(\delta)$  really is an implication (or special case) of the PP, understood as a decision-rule, then PP is deficient, as a decision-rule.

The impossibility result mentioned above is formally proven in the Appendix as Proposition 1. But to put it simply and informally, the reason why one cannot in general simultaneously satisfy WA,  $PP(\delta)$  and Transitivity, is that WA and the aforementioned technical condition will together generate reasoning sequences where the difference in the likelihood of a fatal outcome between consecutive alternatives is not “sufficient”, but where this difference is nevertheless sufficient when it comes to non-consecutive alternatives. Hence, there are situations where WA says that  $X$  is weakly preferred to  $Y$  and  $Y$  is weakly preferred to  $Z$ , yet  $PP(\delta)$  says that  $Z$  is strictly preferred to  $X$ . So, if we add Transitivity to the mix, we get a contradiction.

### 3. IMPOSSIBILITY RESULT FOR THRESHOLD PROVISIO

#### 3.1. Second formulation

Last section demonstrated the tension between the following two ideas: on one hand, the idea that some increases in the likelihood of a catastrophic, or fatal, outcome can be made up for by increasing the likelihood of some desirable outcome in relation to some less desirable, but not catastrophic, outcome; on the other hand, the idea that if one act is sufficiently more likely than another to result in some particular catastrophe, then the second act should be preferred.

---

<sup>21</sup>Why must a decision-rule be transitive? One influential argument is that otherwise, the decision-maker can be *money pumped* (Davidson, McKinsey, & Suppes, 1955). Another argument is that the rule should choose what is better, and betterness is transitive (cf. Broome, on the transitivity of preference, 1991: 118 (en. 20)).

Perhaps the latter idea does not capture the thought behind the Precautionary Principle, however. Maybe a defender of the principle would instead say that other things being equal, if one act confers a relative likelihood on some particular catastrophe that is greater than the likelihood of some chosen outcome (the “threshold outcome”), but another act does not confer a relative likelihood on that same catastrophe that is greater than the threshold, then the second act should be preferred. The idea would be what matters is not how much the two acts differ in terms of how likely they make the catastrophe; what matters is simply that we don’t perform acts that bring with them catastrophic risk that is “too great”, i.e., above some threshold.<sup>22,23</sup> So, a defender of the PP, as a decision-rule, might for instance want to formulate it as follows:<sup>24</sup>

**PP( $\epsilon$ ):** If acts A and B could each result in a fatal outcome, then the latter should be preferred to the former, given that: (i) both fatal outcomes are equally undesirable and (ii) the likelihood that A results in the fatal outcome is above the threshold and (iii) the likelihood that B results in the fatal outcome is not above the threshold.

As is shown at the end of the Appendix, PP( $\epsilon$ ) and PP( $\delta$ ) are logically independent of one another. Hence, one could, for instance, reject PP( $\delta$ ) while accepting PP( $\epsilon$ ). Therefore, as a referee for *Risk Analysis* points out, there is reason to think that, contrary to what Peterson (2006: 599) claims, PP( $\delta$ ) is *not* “so weak that it cannot reasonably be refuted by any advocate of the precautionary principle”. Since a defender of the PP

---

<sup>22</sup>It will be assumed that the likelihood of an outcome can be “not negligibly unlikely”, as referred to in PP( $\delta$ ), without being as likely as the threshold outcome. The idea of “non-negligibility” however plays no special role in this section, and the results discussed below would also hold if we take the threshold to simply distinguish between negligibly and non-negligibly unlikely outcomes.

<sup>23</sup>The idea that risks below a threshold can be ignored is often called the *de minimis* principle. It should be noted that even PP( $\delta$ ) arguably incorporates the *de minimis* principle, since it only requires precautionary measures if the fatal outcomes are “not negligibly unlikely”.

<sup>24</sup>As should be apparent from even the informal statement of PP( $\epsilon$ ), it will be assumed that the threshold is fixed rather than varying. The result I prove holds even if one allows that the threshold can vary somewhat, as long as it is not allowed to vary too much. Formally, one can however avoid the result by allowing the threshold to vary more—e.g. by stipulating that the threshold varies enough to satisfy SAD in every situation! But I doubt that a decision-rule with such a varying threshold can offer practical guidance in different choice situations, be coherent, and non-adhoc.

might favor  $PP(\epsilon)$  over  $PP(\delta)$ , and as the former does not entail the latter, the  $PP(\delta)$  is arguably not weak enough for Peterson's claim to be warranted.

Moreover, there is a reason for favoring  $PP(\epsilon)$  over  $PP(\delta)$ : the former, unlike the latter, can easily accommodate Transitivity. However, a potential reason against  $PP(\epsilon)$ , as compared to  $PP(\delta)$ , is that the former but not the latter directly conflicts with the weak Archimedean condition, WA. To see this, consider a fatal outcome whose likelihood is precisely at the threshold specified by  $PP(\epsilon)$ . Then according to  $PP(\epsilon)$ , *no* increase in the relative likelihood of that outcome could be made up for by increasing the relative likelihood of some desirable outcome. Hence, we have a direct violation of WA. Therefore, the formal analysis of  $PP(\epsilon)$ , in the Appendix, makes use of an Archimedean condition, WA2, that is strictly logically weaker than WA, and which is restricted to acts whose associated fatal outcome is not already precisely at the threshold. Informally, the condition states that:

**Weak Archimedes 2 (WA2):** Some (nonnegligible) increase in the relative likelihood of a fatal outcome can be at least counterbalanced by the relative likelihood of a nonfatal outcome being increased in relation to a strictly worse nonfatal outcome, provided that the likelihood of the fatal outcome was not precisely at the threshold before the increase in likelihood.

### 3.2. Sub-alternative Dominance

Now,  $PP(\delta)$  and WA2 together violate Transitivity, as can be easily verified by replacing WA with WA2 in the proof of Proposition 1 in the Appendix. In contrast,  $PP(\epsilon)$  and WA2 are together consistent with Transitivity (as can also easily be verified). However, taken together,  $PP(\epsilon)$  and WA2 are inconsistent with some very plausible versions of the principle that if a person considers different dimensions of two alternatives,  $P$  and  $Q$ , and she (weakly) prefers alternative  $Q$  to alternative  $P$  on each of these dimensions, then she should (weakly) prefer  $Q$  to  $P$ . Slightly more formally, the principle in question states that:

**Sub-alternative Dominance (SAD):** For any two alternatives  $P$  and  $Q$ , if we can partition the alternatives along some dimensions that the two alternatives have in common, and if it turns out that on each dimension, you weakly prefer the sub-alternative of  $Q$  to the sub-alternative of  $P$ , then you should weakly prefer  $Q$  to  $P$ .

Below I first informally describe “dimensions” and “sub-alternatives”, and discuss an example that illustrates these concepts as well as the requirement of SAD. I then informally describe a condition,  $C$ , that rules out “complementarities” or “interaction effects” between different dimensions, and which needs to hold for SAD to be a requirement of rationality. In the Appendix I however formally state these concepts, as well as SAD and condition  $C$ , and prove (as Proposition 2) that certain instances of SAD—in particular, instances where  $C$  holds—are inconsistent with the combination of  $PP(\epsilon)$  with WA2. The inconsistency arises due to the fact that the catastrophic risk associated with an alternative may fall below the threshold on each dimension but nevertheless fall above the threshold when the alternative is considered as a whole.

As might be apparent, this result needs a technical conditional, similar to the one discussed in the last section, that can be understood as the flip-side of the Archimedean condition. But in this case, the condition ensures that one can (in principle) increase the likelihood of a catastrophe, that is below the threshold stated in  $PP(\epsilon)$ , so much that it crosses the threshold. This version of the condition strikes me as being just as innocent as the one discussed in the last section (and for similar reasons). Hence, I will from now on typically omit explicit references to it, but will instead simply assume it as a background condition.

Informally speaking, the dimensions referred to in SAD are mutually compatible aspects or features of alternatives, that divide alternatives up into non-overlapping parts (the sub-alternatives). Even less formally: dimensions are the sorts of things one typically focuses on when evaluating an alternative (or comparing two alternatives) but finds that evaluating the alternative as a whole is too complex, difficult or demanding.

However, while the dimensions partition an alternative, i.e., they divide the alternative up into non-overlapping and collectively exhaustive parts, the dimensions cannot, for the present purposes (and for reasons that will become apparent), be mutually incompatible in the sense that only one dimension is realized. Instead, dimensions are, informally speaking, different aspects of an alternative, all of which will be realized—but not necessarily at the same time or the same place—if the alternative is chosen.

To take an example, suppose that you are comparing the alternative of living in Boston with that of living in Miami. Since you find it hard to reach an overall judgment about the relative desirability of the two alternatives, you decide to start by comparing them along the different seasons. Now, living in Boston/Miami in summer, say, is an aspect of living in Boston/Miami, and the seasons divide each alternative up into non-overlapping parts. But these parts are not mutually incompatible, since it is possible to, say, live in Miami in the summer and live in Miami in the fall (although not at the same time, of course). So, each season is an example of a dimension, and living in Miami in summer is an example of a sub-alternative of the alternative of living in Miami.

As an illustration of the requirement of SAD, suppose that you prefer living in Boston to living in Miami in any season. Then SAD tells you that you should prefer living in Boston to living in Miami. And in this case, the conclusion suggested by SAD—i.e., to live in Boston—certainly seems supported, if not mandated, by the premises—i.e., by the fact that you prefer living in Boston to living in Miami in any season.

Some examples where we partition an alternative into sub-alternatives cannot, for the present purposes, count as partitioning the alternative along different “dimensions”. An example of such partitioning is familiar from the theory of decision-making under uncertainty, where alternatives are, informally speaking, partitioned along different *events*, i.e., (mutually *incompatible*) sets of states of the world, and rationality is often taken to require that:

**Event-Wise Dominance (EWD):** For any two alternatives  $P$  and  $Q$ , if there is some partition of the state-space into events such that you weakly prefer

$Q$  to  $P$  conditional on every event in that partition, then you should weakly prefer  $Q$  to  $P$ .

EWD will not generate the impossibility proven in the Appendix, since it cannot be the case that for any event  $E_i$ , the unconditional likelihood of some catastrophe is greater than the conditional likelihood of the catastrophe given  $E_i$ .<sup>25</sup> To take a simple (non-catastrophic) example, the unconditional likelihood that it will rain cannot be greater than the conditional likelihood that it will rain given that it is warm and also greater than the conditional likelihood that it will rain given that it is not warm. So, for the present purposes, events are not dimensions. In particular, unlike events, dimensions are not mutually incompatible in the sense that only one dimension is ever realized (as the reader may recall from above). The formal definition of dimensions in the Appendix ensures that they are not mutually incompatible in this sense.

Some instances of SAD are however clearly not requirements of rationality. For instance, suppose the alternatives are personal plans that extend some time into the future, say, one year, and let the dimensions be the different evenings of the year. Now it seems that one could rationally prefer, on each evening, to have a few drinks to not to drink, while preferring to never drink to drinking on every evening; thus violating SAD.

To turn SAD into a general requirement of rationality, that is not undermined by examples like that above, we need to limit it to cases where, informally speaking, the relative likelihoods and desirabilities of outcomes are unaffected when sub-alternatives are combined, in the sense that, say, the desirability or likelihood ranking of two outcomes within a sub-alternative is not reversed when the sub-alternative is chosen along with other sub-alternatives (as compared to when it is chosen without the other sub-alternatives).<sup>26</sup> Thus the formal definition of SAD, in the Appendix, contains a condition,

---

<sup>25</sup>This follows from the assumption that the likelihood is a qualitative probability (except that it may be incomplete).

<sup>26</sup>I thank a referee for encouraging me to discuss this condition in detail, rather than just mentioning it in passing, as I had done in a previous version of this article.

C, that can informally be defined as follows (but which is more formally defined in the Appendix):

**Definition** (C, informal version). *Suppose alternative Z can be partitioned along n dimensions,  $E_1$  to  $E_n$ . Then we say that Z satisfies condition C, with respect to this partition, just in case for any  $E_i$  in the partition, the relative desirability and likelihood of any possible outcome of the sub-alternative of Z in  $E_i$  is independent of whether all of Z is chosen or not.*

Condition C can be illustrated by the following two examples. First, consider again the above drinking example. The decision to drink on some particular night has many possible consequences, one of which being that one feels terrible remorse the next day. Now, suppose that the relative likelihood of this outcome increases as the number of previous nights in which one has taken a drink increases. Then the likelihood independence part of C is violated, and hence SAD (as formally defined) does not apply.

Second, suppose that the undesirability of some possible effect of a chemical  $c$  depends on whether chemical  $c'$  has been used. Now, as first presented, SAD might seem to require that a decision-maker, who is faced with a decision about whether to accept the use of different chemicals, should accept the use of both  $c$  and  $c'$  if she thinks that for each of  $c$  and  $c'$  their use should be accepted. However, due to the supposed negative interaction effect between the two chemicals, the desirabilistic independence part of C is violated. Hence, SAD, as formally defined, does not apply.

As I prove in the Appendix (Proposition 2), the combination of PP( $\epsilon$ ) with WA2 and the aforementioned technical condition violates SAD even when condition C holds. Therefore, PP( $\epsilon$ ) and WA2 together make judgments that, I contend, no good rule for choosing between risky (public or social) policies would make.

### **3.3. The bad of violating SAD**

To take an example, a decision-rule that satisfies both PP( $\epsilon$ ) and WA2 might say that for each part of the economy, the US government is doing enough to prevent environmental

catastrophes caused by activities in that part of the economy, but that the government is nevertheless not doing enough to prevent environmental catastrophes caused by activities in the US economy. (To see this, think of the two dimensions in the table in the Appendix as two parts of the economy.) And this can occur even if condition C holds. That is, the combination of  $PP(\epsilon)$  with WA2 could conflict with SAD, in this context, even if there were no interaction effects between the different parts of the economy.

To take another example, a decision-rule that satisfies both  $PP(\epsilon)$  and WA2 might say that for each chemical in some set of chemicals, we should accept the use of that chemical, but that we should not accept the use of all chemicals in the set. (To see this, think of the two dimensions in the table in the Appendix as two chemicals.) And again, we can construct examples where a decision-rule that satisfies both  $PP(\epsilon)$  and WA2 will make the aforementioned judgment even though condition C holds; for instance, even though there are no negative interaction effects between the chemicals.

Although conflicts like those discussed above between “local” and “global” judgments might not count as inconsistencies, on most senses of the word, they do have the practically unfortunate implication that complex global decision-problems cannot safely be broken down into more manageable units. To take an example, suppose a social planner has to decide whether various chemicals should be allowed in farming. And suppose she uses a decision-rule that satisfies both  $PP(\epsilon)$  and WA2. Then she cannot simply consider each chemical if and when a farmer applies for its approval, but must, once and for all, simultaneously consider all chemicals for which approval might ever be sought. Otherwise, she runs the risk of violating the demands of the PP in her “global” decision problem—i.e., the problem of deciding which set of chemicals to allow—even though she satisfies its demands in each “local” decision—i.e., in each decision about an individual chemical.

Finally, note that a decision-rule that satisfies both  $PP(\delta)$  and WA *also* violates SAD (in addition to violating Transitivity). This is formally proven in the Appendix (as Proposition 3), but informally, the reason is that while the difference in the likelihood

that two alternatives confer on a fatal outcome may be less than “sufficient” (as referred to in  $PP(\delta)$ ) when each dimension is considered on its own, the difference might still be sufficient when all dimensions are considered together. For instance, the difference in the likelihood of a fatal outcome between approving and not approving all chemicals in a set could be sufficient to call for precautionary measures, even though for each chemical in the set, the difference in the likelihood of a fatal outcome between approving and not approving that chemical does not call for such measures.

So, although the results of this paper are negative for both versions of the Precautionary Principle that we have been considering, one might view the results as being more negative for  $PP(\delta)$  than they are for  $PP(\epsilon)$ . For in combination with WA2, the former violates both Transitivity and Sub-alternative Dominance, whereas the latter only violates SAD. On the other hand, recall that  $PP(\epsilon)$ , unlike  $PP(\delta)$ , directly conflicts with WA, which one might see as a reason to favor  $PP(\delta)$  over  $PP(\epsilon)$ .

#### **4. CONCLUDING REMARKS**

The general problem that the results of this paper highlight is that it is hard to square the Precautionary Principle, understood as a rule for choosing in situations of uncertainty, with on one hand the acceptance of any trade-offs between catastrophic risks and chances for more ordinary goods, and on the other hand the idea that some (differences in) catastrophic risks are small enough to overlook.

Most defenders of the Precautionary Principle, as a decision-rule, do want to accept some such trade-offs and also want to ignore some (differences in) catastrophic risks, to avoid the charge of extremism and absolutism. To achieve the latter aim—i.e., ignoring some (differences in) catastrophic risks—it would seem that the PP has to include a proviso stating that some differences in catastrophic risks are small enough to overlook or a proviso postulating some threshold below which catastrophic risk is acceptable. However, since they also want to accept some trade-offs between catastrophic risks

and chances for more ordinary goods, it is hard to see how the first type of proviso could avoid violating Transitivity, and it is hard to see how the second type of proviso could avoid violating Sub-alternative Dominance.<sup>27</sup> In other words, while it is of course possible that defenders of the PP, as a decision-rule, will want to formulate the principle in ways that are different from the formulations considered in this paper, it is hard to see how they can avoid violating at least one condition that, I contend, is satisfied by any good rule for choosing between risky policies.

Another option for those who endorse the Precautionary Principle as a decision-rule for situations of uncertainty, is to simply reject the kind of trade-offs that I have been discussing between catastrophic risks and chances for more ordinary goods. In other words, a defender of the PP might of course simply reject even the weak Archimedean principle discussed above. Whether one can consistently reject Weak Archimedes while avoiding non-absolutism is a question that I will have to leave for another occasion.

Finally, a defender of the Precautionary Principle might argue that the results of this paper have little practical relevance or theoretical importance. For instance, it might be argued that the relevance of these results is undermined by the fact that they concern principles that are limited to situations in which the PP was never intended to be applied. Alternatively, it could be argued that these results have little relevance because they involve certain idealizing assumptions about decision-makers' comparative confidence (or likelihood) judgments. The plausibility of such arguments will also have to be examined on another occasion.

## **ACKNOWLEDGEMENTS**

This paper has benefited from comments and questions that I received when presenting it at the Ethical Risk Management workshop at Texas A&M, the Section for Philosophy at the University of Copenhagen, and the PPE seminar at the Institute for Futures Studies, as well as from discussions I have had with Richard Bradley, Krister Bykvist, Marten

---

<sup>27</sup>With the exception of ad-hoc maneuvers like those mentioned in fn. 24.

Peterson, Daniel Steel, and Katie Steele. Finally, the paper was greatly improved by extremely meticulous comments from three anonymous referees for *Risk Analysis*, as well by very helpful suggestions from the subject editor Terje Aven.

## APPENDIX

To state more formally the conditions and principles that I have been discussing, and to prove the results that I have been referring to, I will make use of a framework very much like the one developed by Peterson (2006). Alternatives are denoted by  $X, Y, Z$ , and are identified with tuples of outcomes, for instance,  $X = [x_1, x_2]$  and  $Y = [y_1, y_2]$  being two alternatives and  $x_1, x_2, y_1$ , and  $y_2$  four possible outcomes.<sup>28,29</sup> Sometimes it will prove useful to denote outcomes by  $a, b, \dots, p$ , where it will be assumed that  $p$  is fatal (i.e., catastrophic), and the desirability ordering corresponds to the alphabetical order ( $a$  being most desirable,  $b$  second most desirable, etc.). The set of all possible outcomes is denoted  $\mathbf{O}$ .

Some more notation:  $x_i \lesssim_d x_j$  means that  $x_j$  is at least as desirable as  $x_i$ ;  $x_i \lesssim_l x_j$  means that  $x_j$  is at least as likely as  $x_i$  (we shall have to assume that both relations are defined for disjunctions and conjunctions of outcomes);  $x^*$  is an outcome that is not negligibly unlikely;  $y^*$  is an outcome that is “sufficiently” more likely than  $x^*$  (as “sufficiently” is used in  $PP(\delta)$ ).  $\lesssim$  is a preference relation on the set of alternatives.  $<$  and  $\sim$  are defined from  $\lesssim$  in the usual way (and similarly for  $\lesssim_l, \sim_l, <_l$ , and  $\lesssim_d, \sim_d, <_d$ ):  $X \sim Y \Leftrightarrow (X \lesssim Y) \& (Y \lesssim X)$ ,  $X < Y \Leftrightarrow (X \lesssim Y) \& \neg (Y \lesssim X)$ .

I have been assuming that the likelihood relation,  $x_i \lesssim_l x_j$ , is a qualitative probability, except that it need not be complete. Formally, this means that I have been assuming that the following three conditions hold for any  $x_i, x_j, x_k \in \mathbf{O}$ :

- If  $x_i \lesssim_l x_j$  and  $x_j \lesssim_l x_k$  then  $x_i \lesssim_l x_k$ .

<sup>28</sup>It is natural to assume that which outcome an alternative results in depends on which state of the world is actual, but states will not be part of the formal discussion.

<sup>29</sup>Unlike Peterson, I don’t order outcomes in alternatives according to their relative likelihood.

- $\emptyset <_l \mathbf{O}$  and  $\emptyset \lesssim_l x_i$ .
- Suppose that  $x_i \& x_j = x_i \& x_k = \emptyset$ . Then  $x_j \lesssim_l x_k \Leftrightarrow x_i \vee x_j \lesssim_l x_i \vee x_k$ .

Following Peterson, I will use the technique of iterating outcomes in alternatives to affect the outcomes' likelihood. For instance,  $a^2$  is  $a$  iterated once, so  $X = [a, b, c]$  becomes  $X' = [a^2, b, c]$  when  $a$  is iterated once. Note that  $a^2$  is meant to denote two instances of outcome  $a$ , such that  $X' = [a^2, b, c]$  is to be read as there being two different ways in which  $X'$  could result in outcome  $a$  (or, given a standard terminology: there are two different "states" under which  $X'$  would result in  $a$ ). So, for instance, if  $a$  is " $n$  lives are lost", then  $a^2$  does not mean that  $2n$  lives are lost, but rather that there are two ways in which the alternative in question could result in  $n$  lives being lost.<sup>30</sup> Therefore, without any additional information, one would expect that  $X'$  is more likely than  $X$  to result in  $n$  lives being lost.<sup>31</sup>

It will be assumed that if there is a bound on how likely one can make an outcome, in relation to other outcomes, by this procedure, then that bound is sufficiently high such that we can use the procedure to make an outcome "sufficiently" more likely (as defined by  $PP(\delta)$ ) than what it previously was, unless the outcome was already (almost) certain to occur.<sup>32</sup> (Similarly, we can use the procedure to take the likelihood over the threshold specified by  $PP(\epsilon)$ .) However, for the below proofs, we do not need to know exactly how much more likely  $X'$  is to result in  $n$  deaths than  $X$ . All we need to assume is that there is some number  $m$  of iterations of  $a$  such that the resulting alternative (with  $m$  iterations of  $a$ ) would be "sufficiently" more likely than  $X$  to result in  $n$  deaths (or has a likelihood above the threshold specified by  $PP(\epsilon)$ ).

I will make the simplifying assumption that differences between at least some relative likelihoods can be ordered. But the results do not depend on it, which is just as well, since the plausibility of the assumption may be questionable in situations of scientific

<sup>30</sup>I thank a referee for making me see the need to add this clarification.

<sup>31</sup>Note that this last claim is true only if we assume that the situation in question is not one of complete ignorance (see Arrow & Hurwicz, 1977).

<sup>32</sup>I thank Krister Bykvist for bringing to my attention the need to make this assumption explicit.

uncertainty. See Stefánsson (2017) for an explanation of how such comparisons can be derived from ordinal comparisons. We define a relation  $\lesssim_{dis}$  on  $\mathbf{O} \times \mathbf{O}$ , and say that  $(a; b) \lesssim_{dis} (c; d)$  whenever the difference in likelihood between  $c$  and  $d$  is at least as great as that between  $a$  and  $b$ . Since differences are symmetric, we have that for any  $x, y \in \mathbf{O}$ :  $(x; y) \sim_{dis} (y; x)$ .

The impossibility results to be proven state that a *decision-rule* cannot simultaneously satisfy certain conditions. The results could be stated in terms of a preference relation, but since the aim is to discuss what rules a social planner should follow when choosing between risky policies, it makes sense to state the results in terms of a decision-rule. A decision-rule, as I am using the term, is a function that takes as input a set of alternatives and a binary (preference) relation on that set, and returns the alternatives that no alternative is preferred to according to the relation, if there are such alternatives (i.e., if the relation is a total or partial order), but otherwise returns the empty set. Hence, a rule is partly defined by the relation on which it is based. For instance, a decision-rule based on expected utility maximizing preferences is different from a decision-rule based on “precautionary preferences”.

Since a decision-rule can be applied to any subset of our set of alternatives, a decision-rule inherits all the features of the underlying preference relation, such as Transitivity or failure thereof. In light of this, it makes sense, as mentioned in the introduction, to understand the PP, on the interpretation considered in this paper, as a *property* of a decision-rule that is based on a preference ordering that satisfies some version of the PP. (But the PP will only be one of many properties of such a rule, since it will also inherit other properties of the underlying preference.)

Here is a formal statement of the principles discussed in section 2:<sup>33</sup>

**Transitivity:** If  $X \lesssim Y$  and  $Y \lesssim Z$  then  $X \lesssim Z$ .

**PP( $\delta$ ):** Let  $X = [x_1, \dots, x_n]$  such that  $x_i \sim_d p$  for precisely one  $x_i$  and  $x^* \lesssim_l x_i$ .

---

<sup>33</sup>It is worth noting that my statement of PP( $\delta$ ) differs from Peterson’s—similarly, my statement of WA differs from Peterson’s statement of A—for instance since, unlike him, I don’t order outcomes in alternatives according to their relative likelihood, and since, unlike him, I employ a four-place difference relation.

Let  $Y = [y_1, \dots, y_n]$  such that  $y_j \sim_d p$  for precisely one  $y_j$  where  $x_i \prec_l y_j$  and  $(x^*; y^*) \lesssim_{dis} (x_i; y_j)$ . Then  $Y \prec X$ .

**Weak Archimedes:** Let  $X = [a, b, \dots, c, p]$ . Then there are some  $j, k, l, m$  such that:<sup>34</sup>

(i)  $X = [a, b, \dots, c, p] \lesssim X' = [a^j, b^k, \dots, c^l, p^m]$ ,

(ii)  $a \prec_l a^j, p \prec_l p^m$ .

We could formulate a stronger Archimedean principle by adding the following condition:  $(x_i \& \neg x_i; x^*) \lesssim_{dis} (p; p^m)$ , to capture the “nonnegligible” clause in the informal statement of WA. However, since the following result holds for all instances of WA, there is no need to work with stronger versions:

**Proposition 1.** *No decision-rule satisfies WA, PP( $\delta$ ), and Transitivity.*

(Note that this result is stronger than Peterson’s Theorem 2 in two respects: first, it uses a weaker Archimedean condition, second, it doesn’t appeal to State-Wise Dominance.)

*Proof.* Let  $X = [a, b, \dots, c, p]$ . Then by WA, there are some  $j, k, l, m$  such that:

(i)  $X = [a, b, \dots, c, p] \lesssim X' = [a^j, b^k, \dots, c^l, p^m]$ ,

(ii)  $a \prec_l a^j, p \prec_l p^m$ .

Now take  $X'$ . Again by WA, there are some  $j', k', l', m', n$  such that:

(i)'  $X' = [a^j, b^k, \dots, c^l, p^m] \lesssim [a^{j+j'}, b^{k+k'}, \dots, c^{l+l'}, p^{m+m'}] = X''$ ,

(ii)'  $a^j \prec_l a^{j+j'}, p^m \prec_l p^{m+m'}$ .

So, by Transitivity,  $X \lesssim X''$ . But, continuing a sequence like this, we will eventually find an alternative  $X'''$ , for which the following holds:

(1)  $X'''$  is weakly preferred to the immediately preceding alternative, so

(2) by Transitivity,  $X \lesssim X'''$ , but

(3) for the fatal outcome in  $X'''$ ,  $p^0: p \prec_l p^0$  and  $(x^*; y^*) \lesssim_{dis} (p; p^0)$ .

So, then by PP( $\delta$ ),  $X''' \prec X$ . Contradiction. □

---

<sup>34</sup>It need not be the case that  $j < k < l < m$  (nor should it be assumed, in the below application of WA, that e.g.  $j = j'$ ).

For the next result, we need some additional notation. Let  $x^* \prec_l z^*$ .  $z^*$  will play the role of the threshold outcome referred to in  $PP(\epsilon)$ .

To ensure that  $PP(\epsilon)$  does not directly contradict the Archimedean condition, we need to weaken it even further:

**Weak Archimedes 2 (WA2):** Let  $X = [a, b, \dots, c, p]$  where  $p \not\prec_l z^*$ . Then there are some  $j, k, l, m$  such that:

(i)  $X = [a, b, \dots, c, p] \preceq X' = [a^j, b^k, \dots, c^l, p^m]$ ,

(ii)  $a \prec_l a^j, p \prec_l p^m$ .

Here is a formal statement of the version of the Precautionary Principles discussed in section 3:

**PP( $\epsilon$ ):** Let  $X = [x_1, \dots, x_n]$  such that  $x_i \sim_d p$  for precisely one  $x_i$  and  $x_i \preceq_l z^*$ .

Let  $Y = [y_1, \dots, y_n]$  such that  $y_j \sim_d p$  for precisely one  $y_j$  and  $z^* \prec_l y_j$ . Then

$Y \prec X$ .

To state Sub-alternative Dominance we need even more notation. Suppose  $E_1, \dots, E_n$  are  $n$  dimensions along which alternatives  $P$  and  $Q$  can be divided into non-overlapping sub-alternatives. Let  $Q(E_i)$  be the sub-alternative of  $Q$  in dimension  $E_i$ . So  $Q(E_i)$  is a tuple of outcomes contained in the tuple identified with  $Q$ . For instance,  $Q = [a, b, c, d]$ ,  $Q(E_1) = [a, b]$  and  $Q(E_2) = [c, d]$ .

To state condition C—which, as the reader may recall, is needed if SAD is to be a general requirement of rationality—we need to introduce the notion of a conditional likelihood, and conditional desirability, given some alternative; and also the notion of conditional likelihood and desirability given some *sub*-alternative. So, let  $x \preceq_l^Z y$  denote that the conditional likelihood of  $y$ , given  $Z$ , is at least as great as the conditional likelihood of  $x$ , given  $Z$ ; and similarly for  $x \preceq_d^Z y$ . To take an example, in the Miami vs. Boston example from section 3, the conditional likelihood statement might be that the

conditional likelihood of skin cancer, given that one lives in Miami, is greater than the conditional likelihood of having to buy a winter jacket, given that one lives in Miami.

In contrast,  $x \lesssim_1^{Z(E_i)} y$  denotes that the conditional likelihood of  $y$ , given the alternative that corresponds to the sub-alternative of  $Z$  in dimension  $E_i$ , is at least as great as the conditional likelihood of  $x$ , given the alternative that corresponds to the sub-alternative of  $Z$  in dimension  $E_i$ ; and similarly for  $x \lesssim_d^{Z(E_i)} y$ . So, in this case, we are conditioning on *just* the alternative that corresponds to the sub-alternative of  $Z$  in dimension  $E_i$ , not the complete alternative  $Z$ . For instance, in the Miami vs. Boston example, one could condition on living in Miami *in summer*, as compared to conditioning on living in Miami. To take an example, the conditional likelihood of a sunburn, given that one lives in Miami in the summer, is greater than the conditional likelihood of suffering from sun deprivation, given that one lives in Miami in the summer.

The notion of conditional likelihood is also useful for stating the assumption that any two dimensions are mutually consistent: We stipulate that for any two  $Q(E_i), Q(E_j)$ ,  $Q(E_i) \& Q(E_j) \sim_1^Q (x_i \& \neg x_i)$ , which, by the assumption of the structure of likelihoods, ensures that the dimensions are mutually consistent.

We can now define the condition needed for SAD to apply as follows (recall that the  $E_i$  are not events):

**Definition** ( $C$ , formal version). *Suppose alternative  $Z$  can be partitioned into  $n$  dimensions,  $\{E_1, \dots, E_n\}$ , such that  $Z = [Z(E_1), \dots, Z(E_n)]$ . Then we say that  $Z$  satisfies condition  $C$ , with respect to the aforementioned partition, just in case for any  $E_i \in \{E_1, \dots, E_n\}$ :*

- For all  $y \in Z(E_i)$  and for all  $x \in \mathbf{O}$  :  $y \lesssim_1^Z x \Leftrightarrow y \lesssim_1^{Z(E_i)} x$ ,
- For all  $y \in Z(E_i)$  and for all  $x \in \mathbf{O}$  :  $y \lesssim_d^Z x \Leftrightarrow y \lesssim_d^{Z(E_i)} x$ .

(Note that the above definition does *not* mean that if  $C$  holds and  $y$  is a *type* of outcome that is contained in, say, both  $Z(E_i)$  and  $Z(E_j)$ , then  $Z$  is just as likely as  $Z(E_i)$  to result in that *type* of outcome. Rather, if we assume that the same type of outcome  $y$  can occur in different sub-alternatives of  $Z$ , say one in each sub-alternative, then  $C$  means

that for any sub-alternative, the relative likelihood of the single *token* occurrence of outcome  $y$  in that sub-alternative is the same as the relative likelihood of the occurrence of the *same* token outcome  $y$  in the complete alternative  $Z$ . This remark also holds for  $\lesssim_d$ .)

Finally, we can now state SAD as a general requirement of rationality:

**Sub-alternative Dominance (SAD):** For any two alternatives,  $P$  and  $Q$ , if there are  $n$  dimensions,  $\{E_1, \dots, E_n\}$ , along which  $P$  and  $Q$  can each be partitioned, such that  $P = [P(E_1), \dots, P(E_n)]$  and  $Q = [Q(E_1), \dots, Q(E_n)]$ , and if both  $P$  and  $Q$  satisfy  $C$  w.r.t. the partition in question, and, finally, if for each  $E_i$ ,  $P(E_i) \lesssim Q(E_i)$ , then  $P \lesssim Q$ .

**Proposition 2.** *No decision-rule satisfies  $PP(\epsilon)$ ,  $WA2$ , and  $SAD$ .*

*Proof.* Let  $X = [a, b, \dots, c, p]$  and  $p \not\prec_1 z^*$ . Then by  $WA2$  there are some  $j, k, l, m$  such that:

- (i)  $X = [a, b, \dots, c, p] \lesssim X' = [a^j, b^k, \dots, c^l, p^m]$ ,
- (ii)  $a \prec_1 a^j, p \prec_1 p^m$ .

	$E$	$\neg E$
$P$	$[a, b, \dots, c, p]$	$[a, b, \dots, c, p]$
$Q$	$[a^j, b^k, \dots, c^l, p^m]$	$[a^j, b^k, \dots, c^l, p^m]$

Table: Violation of SAD

Now consider alternatives  $P$  and  $Q$  in the table, where the sub-alternatives are  $X$  and  $X'$  from above. And assume that  $C$  holds. By  $WA2$ ,  $P(E) \lesssim Q(E)$ ,  $P(\neg E) \lesssim Q(\neg E)$ , so by  $SAD$ ,  $P \lesssim Q$ . But we can construct the example such that, although  $C$  holds,  $z^* \prec_1 p^{m+m}$  (even though  $p^m \lesssim_1 z^*$ ), but  $p^2 \lesssim_1 z^*$ . Then by  $PP(\epsilon)$ ,  $Q < P$ . Contradiction.  $\square$

**Proposition 3.** *No decision-rule satisfies  $PP(\delta)$ ,  $WA$ , and  $SAD$ .*

*Proof.* Let  $X = [a, b, \dots, c, p]$ . Then by  $WA$  there are some  $j, k, l, m$  such that:

- (i)  $X = [a, b, \dots, c, p] \lesssim X' = [a^j, b^k, \dots, c^l, p^m]$ ,
- (ii)  $a \prec_1 a^j, p \prec_1 p^m$ .

Now consider again alternatives  $P$  and  $Q$  in the table. By WA,  $P(E) \lesssim Q(E)$ ,  $P(\neg E) \lesssim Q(\neg E)$ , so by SAD,  $P \lesssim Q$ . But we can construct the example such that although  $C$  holds,  $(x^*; y^*) \lesssim_{dis} (p^2; p^{m+m})$  (even though  $(p; p^m) \prec_{dis} (x^*; y^*)$ ). Then by PP( $\delta$ ),  $Q < P$ . Contradiction.  $\square$

What is the logical relationship between PP( $\delta$ ) and PP( $\epsilon$ )? The answer is that they are logically independent:

PP( $\delta$ ) does not entail PP( $\epsilon$ ) since there are rankings of alternatives that violate the latter but not the former. For instance, if the likelihood of the fatal outcome associated with  $X$  is just above the likelihood of the threshold outcome,  $z^*$ , but the likelihood of the (equally undesirable) fatal outcome associated with  $Y$  is just below the threshold, then PP( $\epsilon$ ) would say that  $Y$  should be preferred to  $X$ , but PP( $\delta$ ) would be consistent with  $X$  being preferred to  $Y$ .

Similarly, PP( $\epsilon$ ) does not entail PP( $\delta$ ) since there are rankings of alternatives that violate the latter but not the former. For instance, suppose that the likelihood of the fatal outcome associated with  $X'$  is greater than the likelihood of the threshold outcome,  $z^*$ , and suppose the same is true of the equally undesirable fatal outcome associated with  $Y'$ . Moreover, suppose that the likelihood of the fatal outcome associated with  $Y'$  is greater than that associated with  $X'$ , and that the difference in likelihood is greater than that between  $x^*$  and  $y^*$ . Then PP( $\delta$ ) would say that  $X'$  should be preferred to  $Y'$ , while PP( $\epsilon$ ) would be consistent with  $Y'$  being preferred to  $X'$ .

## REFERENCES

- Arrow, K. J. and L. Hurwicz (1977). An optimality criterion for decision making under ignorance. In K. J. Arrow and L. Hurwicz (Eds.), *Studies in Resource Allocation Processes*, pp. 463–471. Cambridge University Press.
- Aven, T. (2010). *Misconceptions of Risk*. John Wiley & Sons.

- Aven, T. (2011a). On different types of uncertainties in the context of the precautionary principle. *Risk Analysis* 31(10), 1515–1525.
- Aven, T. (2011b). Response. *Risk Analysis* 31(10), 1538–1542.
- Aven, T. and G. Reniers (2013). How to define and interpret a probability in a risk and safety setting. *Safety Science* 51(1), 223–231.
- Boyer-Kassem, T. (2017a). Is the Precautionary Principle really incoherent? *Risk Analysis* 37(11), 2026–2034.
- Boyer-Kassem, T. (2017b). The Precautionary Principle has not been shown to be incoherent: A reply to Peterson. *Risk Analysis* 37(11), 2039–2040.
- Broome, J. (1991). *Weighing Goods*. Basil Blackwell.
- Chang, R. (2013). Incommensurability (and incomparability). In H. LaFollette (Ed.), *The International Encyclopedia of Ethics*, pp. 2591–2604. Blackwell Publishing Ltd.
- Cox, Jr., L. A. T. (2011). Clarifying types of uncertainty: When are models accurate, and uncertainties small? *Risk Analysis* 31(10), 1530–1533.
- Davidson, D., J. C. C. McKinsey, and P. Suppes (1955). Outlines of a formal theory of value, I. *Philosophy of Science* 22(2), 140–160.
- Gardiner, S. M. (2006). A core Precautionary Principle. *Journal of Political Philosophy* 14(1), 33–60.
- Grant, S. and J. Quiggin (2013). Bounded awareness, heuristics and the Precautionary Principle. *Journal of Economic Behavior & Organization* 93(C), 17–31.
- Jeffrey, R. (1965). *The Logic of Decision*. The University of Chicago Press.
- Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press.

- Karni, E. and M.-L. Vierø (2013). "Reverse Bayesianism": A choice-based theory of growing awareness. *American Economic Review* 103(7), 2790–2810.
- Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971). *Foundations of Measurement*, Volume 1. Academic Press.
- Peterson, M. (2006). The Precautionary Principle is incoherent. *Risk Analysis* 26(3), 595–601.
- Peterson, M. (2007). Should the Precautionary Principle guide our actions or our beliefs? *Journal of Medical Ethics* 33(1), 5–10.
- Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press.
- Peterson, M. (2017a). *The Ethics of Technology. A Geometric Analysis of Five Moral Principles*. Oxford University Press.
- Peterson, M. (2017b). Yes, the Precautionary Principle is incoherent. *Risk Analysis* 37(11), 2035–2038.
- Sandin, P. (1999). Dimensions of the precautionary principle. *Human and Ecological Risk Assessment: An International Journal* 5(5), 889–907.
- Sandin, P., M. Peterson, S. O. Hansson, C. Rudén, and A. Juthe (2002). Five charges against the Precautionary Principle. *Journal of Risk Research* 5(4), 287–299.
- Savage, L. (1954). *The Foundations of Statistics*. John Wiley & Sons.
- Steel, D. (2014). *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy*. Cambridge University Press.
- Steele, K. (2006). The Precautionary Principle: a new approach to public decision-making? *Law, Probability and Risk* 5(1), 19–31.
- Stefánsson, H. O. (2017). What is 'real' in Probabilism? *Australasian Journal of Philosophy* 95(3), 573–587.

- Steup, M. (2018). Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 ed.). Metaphysics Research Lab, Stanford University.
- Temkin, L. S. (2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford University Press.
- Tversky, A. and D. Kahneman (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.*, pp. 84–100. Cambridge University Press.
- Vlek, C. (2011). Straightening out the grounds for precaution: A commentary and some suggestions about Terje Aven's "On different types of uncertainties". *Risk Analysis* 31(10), 1534–1537.
- Walker, O. and S. Dietz (2011). A representation result for choice under conscious unawareness. *Grantham Institute Working Paper*.
- Warner North, D. (2011). Uncertainties, precaution, and science: Focus on the state of knowledge and how it may change. *Risk Analysis* 31(10), 1526–1529.