How Many Alternatives? Partitions Pose Problems for Predictions and Diagnoses
Michael Smithson

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

How Many Alternatives? Partitions Pose Problems for Predictions and Diagnoses

Michael Smithson

This paper focuses on one matter that poses a problem for both human judges and standard probability frameworks, namely the assumption of a unique (privileged) and complete partition of the state-space of possible events. This is tantamount to assuming that we know all possible outcomes or alternatives in advance of making a decision, but it is clear that there are many practical situations in prediction, diagnosis, and decision-making where such partitions are contestable and/or incomplete. The paper begins by surveying the impact of partitions on the choice of priors in formal probabilistic updating frameworks, and on human subjective probability judgements. That material is followed by an overview of strategies for dealing with partition dependence, including considerations of how a rational agent's preferences may determine the choice of a partition.

Introduction

Many quantitative sciences and professions based on them adhere to canons of statistical inference (e.g. confidence intervals) and probabilistic reasoning (e.g. Bayesian updating) when it comes to induction, prediction, and diagnosis. However, scientists, practitioners and other experts are human beings, and there is a vast empirical literature demonstrating that humans reason under uncertainty in ways that depart from the prescriptions of probability theory. Moreover, the prescriptions themselves have important limitations.

Michael Smithson is Professor in the Psychology Department at The Australian National University. Correspondence to: Dr Michael Smithson, The Australian National University, Psychology, Bldg 39, Canberra 0200, Australia. Email: michael.smithson@anu.edu.au

This paper focuses on one matter that poses a problem for both human judges and standard probability frameworks; namely, the assumption of a unique (privileged) and complete partition of the state-space of possible events. This assumption amounts to believing that we know all possible outcomes or alternatives in advance of making a decision. We begin by considering arguments that both human judgements and standard probability frameworks are strongly dependent on the choice of a partition, and establishing that there are many practical situations in prediction, diagnosis, and decision-making where such partitions are contestable and/or incomplete.

This material is followed by an overview of strategies for dealing with partition dependence. These include attempts to construct probabilistic frameworks that are not partition dependent, and recent studies indicating that human judges utilizing tools from one such framework nonetheless remain strongly partition dependent. We also investigate criteria that a rational agent may apply in choosing among alternative permissible partitions. This choice turns out to hinge on the compatibility between the agent's utilities regarding correct versus incorrect choices and threshold probabilities that determine when one alternative is chosen over another.

The Road to Partition: Normative and Descriptive Issues

Partition Dependence in Probability Theory

Standard probability theories applied to settings in which there is a finite set of possible events are partition dependent. On grounds of insufficient reason, a probability of $1/K$ is assigned to K mutually exclusive possible events when nothing is known about the likelihood of those events. For example, in a race involving three greyhounds, a decision-maker who knows nothing about any of the dogs would assign a value of 1/3 to the probability of each greyhound winning. The intuitive appeal of an *ignorance prior* was captured by early probability theorists such as Laplace and Leibniz in the form of the "principle of insufficient reason", whereby if there is no obvious reason why one event should happen rather than another event, then both events should be treated as equally probable. Moreover, even under alternative probability assignments the probabilities of the K events must sum to one, meaning that the entire probability mass is concentrated on that set of events.

However, it is not difficult to find scenarios in which the choice of K is problematic. Indeed there are many practical situations in which no privileged partition exists and yet the choice of partition crucially affects decision-making. The web site of the Centre for Evidence-Based Medicine (2007) at the University of Toronto presents an example of a female patient who has been administered a serum ferritin test for diagnosing iron-deficiency anaemia. Her test yields 40 mmol/l. Their hypothetical evidence for the diagnostic application of this test is shown in the upper part of Table 1, suggesting that a result of 45 mmol/l or less has a likelihood ratio of 8.24 and a post-test probability of $70/85 = 0.82$ of having the disorder. The obvious recommendation for this patient is treatment.

Table 1 Hypothetical Anaemia Test Scenario

Serum ferritin test result	Iron-deficiency anaemia		Likelihood ratio
	Present	Absent	
First test scenario			
Positive (≤45 mmol/l)	70	15	8.24
Negative (>45 mmol/l)	15	135	0.20
Total	85	150	
Second test scenario			
≤35 mmol/l	60	3	35.29
>35 to ≤75 mmol/l	15	35	0.76
>75 mmol/l	10	112	0.16
Total	85	150	

However, what if the partition used for the serum ferritin test was the one shown in the lower part of Table 1? The evidence base is the same, but now the patient lands in a category whose likelihood ratio is only 0.76, hardly sufficient to compel us to recommend treatment. Perhaps further investigation or tests would be warranted instead. The point is that neither partition is normatively preferable to the other according to probability theory, but each yields a different decision. In this situation, the choice between these (or other alternative) partitions must rest on considerations outside of probability theory.

For a somewhat more generic type of example, following Walley (1996), consider a bag containing 10,000 variously coloured marbles whose colours are completely unknown to us. How should we use the principle of insufficient reason to judge the probability of drawing a red marble from this bag? We could partition colours into red versus not red, thereby invoking the ignorance prior probability $P(\text{red}) = 1/2$, but we could just as reasonably partition colours in a more fine-grained way (e.g. {red, other warm colours, cool colours, white, grey, black}) and thereby induce a smaller prior probability for $P(\text{red})$. We could, for that matter, use a partition comprising thousands of colours. The ignorance prior assignment to $P(\text{red})$ can be made arbitrarily small, depending on how finely we partition colours.

When partitions are indeterminate, partition dependence poses a difficult problem for subjective probability assignments. Note that this is not the same problem as unknown probabilities over a unique and complete partition (e.g. where we know that there are only red and black marbles in the bag but do not know how many of each). It is much more profound. Bayesian theory falls silent here, as does any other standard probability theory. In the absence of a uniquely privileged partition, there is no defensible prior probability assignment to be made. Nonetheless, it should be apparent that partition dependence is a greater problem for the Frequentist than for the Bayesian framework. Bayesians may or may not assign priors using a sample space, whereas Frequentists always must work with a sample space.

Two separable problems for partitions are contained in the marble example. One is an incomplete account of possible events. A privileged partition may be attainable in principle, but we lack the necessary information. The other problem is the absence of a privileged partition even when one has a complete account of those possibilities. Shafer (1976) presented an example of this problem as a motivation for the belief functions framework. He asked whether the probability of life existing in a newly discovered solar system should be partitioned into {life, no life} or {life, planets without life, no planets}. This difficulty arises naturally when a decision must be made that involves a threshold or interval on a continuum, as in the serum ferritin test scenario. We shall revisit this particular problem later.

Clearly these kinds of partition problems can and do occur in science and practice. A typical example where incomplete partitions arise is a clinical trial in which a new treatment is compared with an established one and/or a placebo condition. The source of incompleteness in the partition is the possibility of side-effects from the new treatment. We may have no good reason to presuppose what the variety of possible side-effects might be, and any partitioning of such effects in the tradition of probability theories would have to resort to fiat and the use of a "catch-all" category for unforeseeable effects.

Partition Dependence in Human Judges

Human judges are strongly influenced by partitions in their intuitive probability assignments. While some aspects of that influence are reasonable according to probability theories, others are not. Two important manifestations of partition dependence are distorted judgements of likelihoods of compound events and anchoring on an ignorance prior.

Turning first to the compound event likelihood issue, a seminal study by Fischhoff, Slovic, and Lichtenstein (1978) concerning people's assignments of probabilities to possible causes of a given outcome (e.g. an automobile that will not start) revealed that possible causes that were explicitly listed received higher probabilities than when the same causes were implicitly incorporated into a "Catch-All" category of additional causes. One explanation proposed for this effect amounts to the old (and debatable) proverb "Out of sight, out of mind", and the effect has since been referred to as the "Catch-All Underestimation Bias" (CAUB) and also sometimes the "pruning bias" (Russo and Kolzow 1994).

At least three explanations have been proposed for the CAUB:

1. Unlisted causes are not as available to a person's mental representation of the situation, and therefore not rated as highly likely.
2. People may perceive ambiguity in a list that is incomplete and mentally redefine some of the items on the list by adding other unlisted causes to them (Hirt and Castellan 1988).
3. A list that is incomplete may be perceived as lacking credibility, so people inflate the probabilities of the explicitly listed causes (Dubé-Rioux and Russo 1988).

Russo and Kolzow (1994) conducted further studies and found the most evidence for the unavailability explanation and the least evidence for the credibility explanation. However, Bonini and Caverni (1995) provided evidence from the literature and their own experiments that casts doubt on all three explanations. For instance, they found that making the unlisted causes more available to people did not decrease the CAUB.

In a more general vein, Support Theory (Rottenstreich and Tversky 1997; Tversky and Koehler 1994) is a framework that begins with the claim that people do not follow the logic of conventional probability theory. Instead, unpacking a compound event (e.g. cancer) into disjoint components (e.g. breast cancer, lung cancer, etc.) tends to increase the perceived likelihood of that event. An immediate implication is that unpacking a hypothesis and/or repacking its complement will increase the judged likelihood of that hypothesis. Human judges are not behaving according to probability theory here, but what are they doing instead? One explanation is that they intuitively evaluate the likelihood of a compound event by the number of ways they think that it can happen. The more distinct paths to an outcome they are aware of, the more likely that outcome seems.

Rottenstreich and Tversky (1997) also found that although unpacking a category increases its subjective probability, separate assessment of the unpacked subcategories increases the total subjective probability even more. Sloman et al.'s (2004) review of investigations into this phenomenon concluded that the separate evaluation effect is stronger than the unpacking effect.

Partly as an outgrowth of this latter finding, Fox and Rottenstreich (2003) conjectured and demonstrated that subjective probability judgements are typically biased towards the ignorance prior determined by the partition salient to the judge. That is, people anchor on a uniform distribution of $1/K$ across all K possible events, even when taking into account prior evidence of how likely each event is. Because those adjustments typically are insufficient (Tversky and Kahneman 1974), judges' intuitive probability assignments are biased toward probabilities of $1/K$. Fox and Rottenstreich demonstrated partition dependence by experimentally manipulating which partition was made salient to judges and showing that judges' probability judgements were influenced accordingly.

Partition dependence poses problems for probability judgements in two respects. First, it may be *unjustified* because there is an arguably "correct" partition. For instance, Fox and Rottenstreich pose the question of how likely Sunday is to be the hottest day of the week. The principle of insufficient reason would suggest that 1/7 is the correct answer, so their demonstration that people can be induced to partition the events into just two possibilities (by being primed with the idea that Sunday either is or is not the hottest day) and therefore arrive at 1/2 indicates that those people are miscued into an incorrect partition. However, as we have seen, a more profound difficulty is that partitions may be *arbitrary*.

Fox and Clemen (2005) link partition dependence to the CAUB and unpacking–repacking phenomena described above, and argue that anchoring to the ignorance prior dictated by one's partition accounts for these phenomena. They find evidence of

partition dependence for continuous as well as discrete events and continuous variables, demonstrate that domain expertise can reduce but does not eliminate the anchoring effect, and show that even respondents highly trained in formal decision theory are susceptible to it.

Fox and Levav (2004) point out that the consequences of partition dependence extend to judgements of conditional probabilities as well. The following illustration parallels an example from their paper. Three pharmaceutical companies have been developing a new class of cancer-fighting drugs (A, B, and C). We have just learned that an independent laboratory compared the effectiveness of A and C, finding definitively that A is more effective than C. What is the probability that a study comparing all three drugs will conclude that A is more effective than both B and C? If we had to choose between administering drugs A or B on the basis of the current evidence, should we prefer A or not?

According to Fox and Levav's empirical findings, many of us will reason that C has been eliminated as the strongest drug, so that leaves two equally likely possibilities: A or B. Therefore, the probability that A is the strongest of the three is 1/2 and there is no reason to prefer A over B. The correct answer may be arrived at by considering which of the possible orderings among the drugs have been eliminated. The possible orderings are {ABC, ACB, BAC, BCA, CAB, CBA}, wherein the sequence of the letters indicates their ranking (e.g. ABC indicates that A > B > C). The latter three orderings have been ruled out because they are the only ones in which C outranks A. Drug A is the strongest in two out of the first three orderings, so the probability that a study comparing all three drugs will conclude that A is more effective than both B and C is 2/3. We should definitely prefer A over B. Fox and Levav argue that the heuristics attracting people to the conclusion that the probability is 1/2 involve "editing" the partition from three events to two on the basis that C has been "ruled out."

Dealing with Partition Dependence

Having established that partition dependence presents difficulties for human judges and, under some conditions, for standard probability frameworks, we now review proposals for dealing with those difficulties. A natural classification of these proposals distinguishes those that apply when a privileged or at least agreed-upon partition is attainable from those that apply when it is not:

1. Where a privileged or agreed-upon partition is attainable:

   - Debiasing strategies.
   - Establishing criteria for choosing partitions.

2. Where there is no privileged or agreed-upon partition:

   - Using diverse partitions.
   - Modelling partition-dependence effects.
   - Using (non-standard) probabilistic frameworks that avoid partition dependence.

While it is not possible to cover all of these possibilities in depth, a general overview and examples of each will be provided here.

Debiasing Strategies

So-called "debiasing" procedures presume a privileged partition (whether on normative or other grounds). They are not particularly interesting for our purposes here, so we shall review them only briefly. Fox and Clemen (2005) build on Russo and Kolzow's (1994) recommendations for avoiding the CAUB. First, judges must be directed to fully and even-handedly elaborate the partition itself, unpacking all subcategories. Second, they must systematically evaluate support for each subcategory (Fox and Clemen do not make specific recommendations but there is a large literature on methods for eliciting subjective probability judgements).

Finally, along with most normative prescriptions, Fox and Clemen admonish judges to map their impressions of relative support for each subcategory onto a set of numbers that obey the rules of probability theory. They caution that the tendency of judges to anchor on the ignorance prior may be the most difficult bias to combat because it is least available to conscious reflection. More recently, See, Fox, and Rottenstreich (2006) have investigated the effects of partition dependence on human judges' updating of probability judgements as they learn new evidence.

Criteria for Choosing Partitions

Where there are no normatively privileged partitions, partitions still may of course be chosen or constructed on other grounds. In medical practice, for example, it could be argued that the variety of alternatives available to the practitioner is driven by technical and legal contingencies. Nevertheless, any partition established by fiat can be critically examined for consistency with a decision-maker's preferences. The potential for using such preferences to adjudicate between alternative partitions appears not to be widely known, however, and so we shall devote some space to exploring that potential here.

In the context of legal standards of proof, Connolly (1987) points out that a typical threshold probability of guilt associated with the phrase "beyond reasonable doubt" is in the [0.9, 1] range. For a logically consistent juror, a threshold probability of 0.9 implies the difference between the utility of acquitting versus convicting the innocent is nine times the difference in the utility of convicting versus acquitting the guilty. Note that this is the only restriction imposed by such a threshold—although as DeKay (1996) points out, it is not equivalent to a Blackstone-like 1-to-10 ratio of erroneous convictions versus erroneous acquittals.

Connolly demonstrates that the utility assignments to the four possible outcomes (convicting the guilty, acquitting the innocent, convicting the innocent, and acquitting the guilty) that are compatible with such a high threshold probability are counter-intuitive. Specifically, "… if one does [want to have a threshold of 0.9], one must be prepared to hold the acquittal of the guilty as highly desirable, at least in

comparison to the other available outcomes" (Connolly 1987, 111). He also shows that more intuitively reasonable utilities lead to unacceptably low threshold probability values.

Smithson (2006) shows that the incorporation of a third middle option (such as the Scottish Not Proven verdict) with a suitable threshold can resolve this quandary, permitting a rational (subjective expected utility) agent to retain a high conviction threshold and still regard false acquittals as negatively as false convictions. The price paid for this solution is a more stringent "standard of proof" for outright acquittal.

In a series of empirical studies of mock-juror decision-making in which a Not Proven alternative was made available, Smithson, Gracik, and Deady (2007) found the judged probability of guilt for those participants returning a Not Proven verdict was firmly in the middle range of the scale, significantly higher than the mean for those returning Not Guilty and lower than the mean for those returning a Guilty verdict. In fact, the Not Proven probabilities were within the range where a rational agent utilizing a high threshold for conviction would return a Not Proven verdict.

The main point here is that a consideration of preferences as expressed by relative positions of utilities can aid in the choice of a partition. An illustration in medical diagnostic classification is an actuarial violence risk assessment tool presented in Steadman et al. (2000), using discharged hospital patients' follow-up data. Commission of a violent act during the first 20 weeks following hospital discharge was coded as a dichotomous outcome. Of the sample of 939 discharged patients included for study, 18.7% committed a violent act within 20 weeks after discharge.

Without any apparent justification, Steadman et al. classify any case with a predicted probability of violence greater than twice the base prevalence rate (>0.37) in the "high-risk" category, and any case whose predicted probability of violence is less than one-half the base prevalence rate (<0.09) in the "low-risk" category. The middle group is "unclassified". They therefore have two probability thresholds separating three risk-classifications (i.e. a three-fold partition). Smithson (2006) demonstrates that this partition and its thresholds correspond to the preferences of a rational agent for whom the consequences of wrongly classifying a violent patient are weighted much more heavily than the consequences of wrongly classifying a non-violent patient.

When would a rational decision-maker actually require a three-fold rather than a two-fold partition? Smithson (2006) presents an answer in terms of compatibility between preferences regarding the thresholds and preferences regarding utility differences. Suppose we have two possible states of interest: A = diseased and ~A = not diseased. Table 2 shows the utilities given A and ~A, associated with a two-fold and three-fold partition of acts R_j. We assume $H_j > H_{j-1}$ and $G_{j-1} > G_j$ for any j. A straightforward argument shows that if the odds of A, $w_A > w_{j-1j} = (G_{j-1} - G_j)/(H_j - H_{j-1})$, then the decision-maker should prefer act R_j over R_{j-1}. The threshold odds w_{j-1j} therefore are determined by the ratio of utility differences.

Now, suppose in the two-fold partition that the decision-maker stipulates $(G_0 - G_1)/(H_1 - H_0) < 1/3$ but also insists that w_{01} is $\geq 1/2$. That is, the decision-maker holds that

Table 2 Utilities for Two-fold and Three-fold Partitions of Acts

Reality	Act		
	R_1 = Treatment	R_0 = Send home	
Two-fold partition			
A = Diseased	H_1	H_0	
~A = Clear	G_1	G_0	
	R_2 = Treatment	R_1 = Further tests	R_0 = Send home
Three-fold partition			
A = Diseased	H_2	H_1	H_0
~A = Clear	G_2	G_1	G_0

the difference between being right and wrong about a diseased patient is at least three times greater than that between being right and wrong about a patient with no disease, but also that he/she is not willing to use an odds of less than 1/2 to assign a patient to the Treatment category. These preferences cannot be accommodated in a two-fold partition but can be accommodated in a three-fold one, for reasons illustrated in Figure 1. The diagrams in Figure 1 are a geometric interpretation of the restrictions placed by threshold odds and utility differences on each other. Given the differences $G_0 - G_1$ and $H_1 - H_0$ in the two-fold partition diagram, the slope of the line connecting G_0 and H_1 is $-w_{01}$, the negative of the threshold odds.

The left-hand diagram shows that no line corresponding to $w_{01} \geq 1/2$ can connect G_0 and H_1 anywhere within the region dictated by $(G_0 - G_1)/(H_1 - H_0) = w_{01} < 1/3$, simply because $w_{01} \geq 1/2$ and $w_{01} < 1/3$ cannot hold simultaneously. However, the right-hand diagram shows by example that the introduction of a middle option allows both $w_{12} \geq 1/2$ and $(G_0 - G_2)/(H_2 - H_0) < 1/3$ (see Smithson [2006] for a discussion of the properties and uses of this kind of diagram). More generally, a $K + M$ alternative partition can resolve incompatibilities between utility and probability threshold assignments in a partition with K alternatives. Thus, preferences regarding differences among utilities and probability thresholds can guide one's choice of partition.
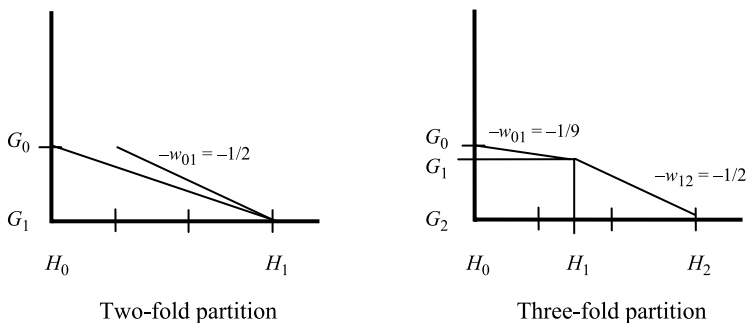

Two-fold partition　　　　　Three-fold partition

Figure 1 Utilities and Threshold Odds.

Diverse Partitions and Modelling Partition-dependence Effects

The strategies for dealing with partition dependence when no partition can be agreed on are less well-established than those that apply when a unique partition can be found. An obvious suggestion is to use more than one partition and compare the resulting probability assignments for inconsistencies due to partition dependence. A somewhat more sophisticated but untested possibility (Fox and Clemen 2005) is to model the influence of partitions on judges' assignments and then remove the estimated influence, thereby "debiasing" the subjective probabilities. If subjective probabilities anchored on ignorance priors can be modelled as convex sums of the appropriate ignorance prior and an evidence-driven component, and if the weighting given to the ignorance prior is constant, then by obtaining a judge's probability assignments under a variety of partitions one could estimate and isolate the evidence-driven component.

Convex-sum mixture models of this kind have been proposed but remain relatively untested (see Clemen and Ulu 2008; Fox and Rottenstreich 2003; See, Fox, and Rottenstreich 2006). Moreover, a recent literature has emerged in other contexts concerning appropriate statistical techniques for modelling continuous proportions such as subjective probabilities (see Smithson and Verkuilen 2006). However, rather strong assumptions are required for these approaches to be effective, and considerable empirical work and further theoretical development would be needed to test those assumptions.

Partition-independent Probability Frameworks

Perhaps the most radical strategy is to construct a probability framework that is not partition dependent. Efforts to do so have appeared in the growing literature on generalized probability theories, and also in behavioural economics (for example, Grant and Quiggin 2004). Walley (1991, 1996) argues on normative grounds that while partition dependence poses a problem for likelihood judgements constrained to a single probability, imprecise probability frameworks can avoid it entirely. He proposes that when judges are permitted to provide a lower and upper probability judgement (i.e. imprecise probabilities), every ignorance prior should consist of vacuous probabilities {0,1}. For example, someone who knows nothing about the form of two players at the start of a chess match may not prefer any probability in the interval from zero to one when assessing the probability that one of the players will win. They could assign a lower probability of zero and an upper probability of one regardless of whether the partition in question is {win, not win} or {win, draw, lose}. If forced to use just one number they are compelled to become partition dependent, but if they can use lower and upper probabilities then they need not be partition dependent in their assignments.

Moreover, imprecise probability judgements that are influenced by this ignorance prior should not depend on the partition. Thus, *a posteriori* imprecise likelihood judgements influenced by this ignorance prior also should be partition independent. Walley

(1996) calls this the Representation Invariance Principle. He proposes an updating method (the Imprecise Dirichlet Model) that is partition independent and has generated interest within the community of imprecise probability theorists.

The question of whether imprecise probability judgements by humans exhibit partition dependence is important. If imprecise probability judgements are not strongly influenced by partitions, then imprecise subjective probabilities might be preferable to precise subjective probabilities under conditions where partition dependence is problematic. On the other hand, if subjective imprecise probability judgements are influenced by partitions then it still may be possible to train judges to make judgements in line with Walley's proposals.

There has been little investigation of partition priming effects on imprecise probability judgements. Smithson, Bartos, and Takemura (2000) investigated whether mathematically naïve participants would adhere to Walley's principles under what they called "sample space ignorance" (i.e. incomplete partitions). They found that in both Australian and Japanese samples, a majority of participants gave lower and upper probabilities of {0,1} under complete sample-space ignorance with no partition information. They also found that when evidence concerning the likelihood of events was presented to participants, they narrowed the gap between their lower and upper probability estimates, albeit not as rapidly as Walley's updating methods would prescribe. However, they also found that the gap was larger for events that had not yet been observed than for those that already had been observed at least once. This finding suggests that people may not treat strength and weight of evidence independently even when given the means to do so.

Smithson et al. (2000) also investigated the impact of partition cues from evidence and from the event whose likelihood was being judged. They found effects when one event was perceived as including another (e.g. the likelihood of a marble of any heretofore unobserved colour versus a heretofore unobserved orange marble being drawn from a bag of marbles).

These findings suggest that imprecise probability judgements might be influenced by partition cues. Smithson and Segale (2009) replicated and extended several experiments of Fox and his colleagues, and found that naïve judges are just as strongly influenced by partitions when making imprecise probability judgements as they are when making precise probability judgements. Moreover, they demonstrated that many judges anchor on $1/K$ as the midpoint of their lower and upper probability judgements.

Figure 2 shows the distributions of the midpoints in imprecise probability intervals given in response to Fox and Rottenstreich's (2003) question of how likely Sunday is to be the hottest day of the week, under their two experimental conditions. The "case" prime makes a two-fold partition salient (either Sunday will or will not be the hottest day), whereas the "class" prime makes a seven-fold partition salient (Sunday is one out of seven days in the week). Smithson and Segale demonstrated via mixture models that most judges are shifting between a probability interval whose midpoint is close to 1/2 and one whose midpoint is close to 1/7, depending on the prime.

They also found that judges simultaneously influenced by rival partitions assigned lower and upper probabilities whose interval encompassed the rival priors entailed by
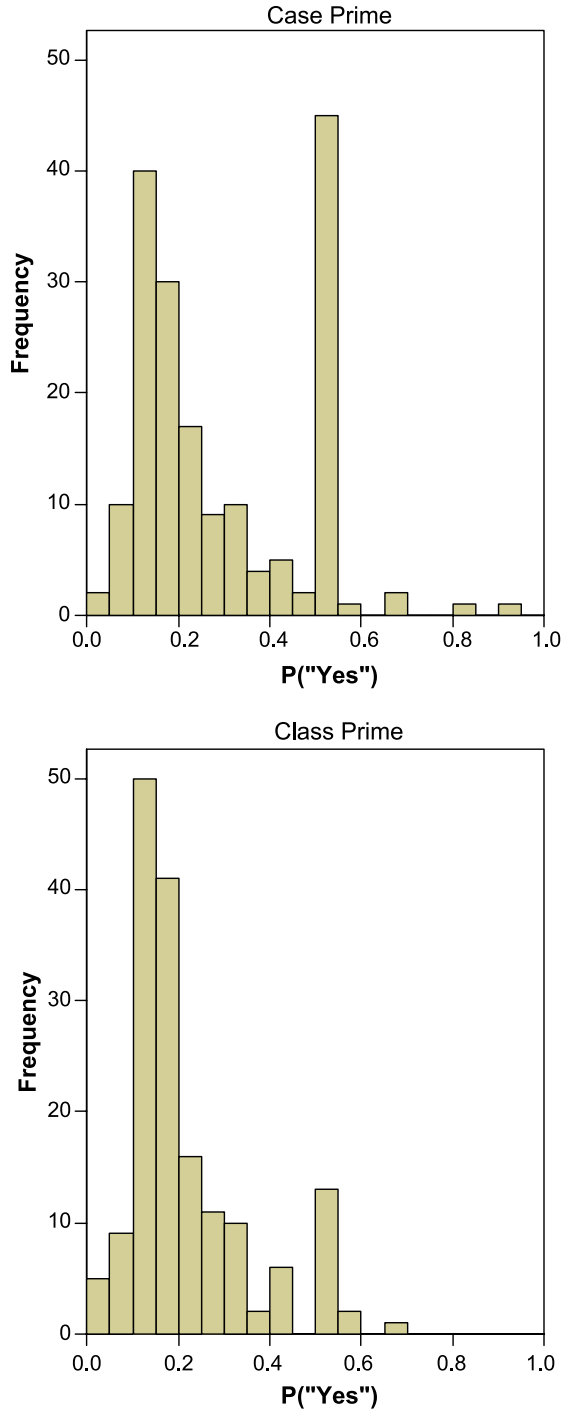
Figure 2   Weather Task *P*("Yes") Distributions.

those partitions. In the hottest-day task, many judges given the "case" prime provided a probability interval that included both 1/7 and 1/2. They posit that this tendency will be most pronounced when there is a privileged or preferred partition but an alternative partition is made salient to the judge. The findings of Smithson and Segale are sobering, but the possibility still remains that judges may be trained to make imprecise probability assignments that are not partition dependent.

Concluding Remarks

We began with an argument that partition dependence poses problems for standard probability theories when a complete, normatively privileged partition is unavailable, and for human judges even when such a partition is available. It should be evident by now that the conditions that make partition dependence problematic are neither trivial nor uncommon. In fact, they apply to many of the real-world decision-making contexts that typify the application of frameworks such as evidence-based medicine. More often than not, we do not have a complete listing of possible outcomes or alternatives. More often than not, partitions are contestable and socially constructed.

The survey of currently available strategies for combating the effects of partition dependence reveals a rather meagre repertoire and suggests there is much to be done. Little is known about how conventional partitions in decisional contexts come to be established, or even when particular partitions are likely to be salient to practitioners or other experts. Likewise, very little work has been done on the issue of appropriate criteria for selecting or constructing partitions. Finally, while emerging frameworks in generalized probability theory that are partition independent are promising, fully-fledged frameworks with well-developed statistical methods still seem somewhat far off. And the question of whether human judges can use those frameworks to shed the habit of partition-dependent subjective probability judgements remains unanswered. At least some aspects of the work to be done will necessarily be domain specific, so there is ample opportunity here for scientists and practitioners to tailor strategies that will be appropriate to their domains.

References

Bonini, N., and J.-P. Caverni. 1995. The "catch-all underestimation bias": Availability hypothesis vs. category redefinition hypothesis. *Current Psychology of Cognition* 14 (33): 301–22.

Centre for Evidence-Based Medicine. 2007. University of Toronto. Available from http://www.cebm.utoronto.ca/practise/ca/diagnosis/important.htm; INTERNET.

Clemen, R. T., and C. Ulu. 2008. Interior additivity and subjective probability assessment of continuous variables. *Management Science* 54 (4): 835–51.

Connolly, T. 1987. Decision theory, reasonable doubt, and the utility of erroneous acquittals. *Law and Human Behavior* 11 (2): 101–12.

DeKay, M. L. 1996. The difference between Blackstone-like error ratios and probabilistic standards of proof. *Law and Social Inquiry* 21 (1): 95–132.

Dubé-Rioux, L., and J. E. Russo. 1988. An availability bias in professional judgment. *Journal of Behavioral Decision Making* 1 (4): 223–37.

Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception Performance* 4 (2): 330–44.

Fox, C. R., and R. T. Clemen. 2005. Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science* 51 (9): 1417–32.

Fox, C. R., and J. Levav. 2004. Partition-edit-count: Naïve extensional reasoning in judgment of conditional probabilities. *Journal of Experimental Psychology: General* 133 (4): 626–42.

Fox, C. R., and Y. Rottenstreich. 2003. Partition priming in judgment under uncertainty. *Psychological Science* 13 (3): 195–200.

Grant, S., and J. Quiggin. 2004. Conjectures, refutations and discoveries: incorporating new knowledge in models of belief and decision under uncertainty. Paper presented at the 11th International Conference on the Foundations and Applications of Utility, Risk and Decision Theory (FUR XI—Paris), under the joint auspices of the Ecole Nationale Supérieure d'Arts et Métiers (ENSAM) and the Ecole Spéciale des Travaux Publics (ESTP), Paris, 2 July.

Hirt, E. R., and N. J. Castellan, Jr. 1988. Probability and category identification in the fault tree paradigm. *Journal of Experimental Psychology: Human Perception Performance* 14 (1): 112–31.

Rottenstreich, Y., and A. Tversky. 1997. Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review* 104 (2): 406–15.

Russo, J. E., and K. J. Kolzow. 1994. Where is the fault in fault trees? *Journal of Experimental Psychology: Human Perception Performance* 20 (1): 17–32.

See, K. E., C. R. Fox, and Y. S. Rottenstreich. 2006. Between ignorance and truth: Partition dependence and learning in judgment under uncertainty. *Journal of Experimental Psychology: Learning, Memory Cognition* 32 (6): 1385–402.

Shafer, G. 1976. *A mathematical theory of evidence.* Princeton, NJ: Princeton University Press.

Sloman, S., Y. Rottenstreich, E. Wisniewski, C. Hadjichristidis, and C. R. Fox. 2004. Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory Cognition* 30 (3): 573–82.

Smithson, M. 2006. Scale construction from a decisional viewpoint. *Minds and Machines* 16 (3): 339–64.

Smithson, M., T. G. Bartos, and K. Takemura. 2000. Human judgment under sample space ignorance. *Risk Decision and Policy* 5 (2): 135–50.

Smithson, M., L. Gracik, and S. Deady. 2007. Guilty, not guilty, or …? Multiple verdict options in jury verdict choices. *Journal of Behavioral Decision Making* 20 (5): 481–98.

Smithson, M., and C. Segale. 2009. Partition priming in judgments of imprecise probabilities. *Journal of Statistical Theory and Practice* 3 (1): 169–82.

Smithson, M., and J. Verkuilen. 2006. A better lemon-squeezer? Maximum likelihood regression with beta-distributed dependent variables. *Psychological Methods* 11 (1): 54–71.

Steadman, H. J., E. Silver, J. Monahan, P. S. Appelbaum, P. Clark Robbins, E. P. Mulvey, T. Grisso, L. H. Roth, and S. Banks. 2000. A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior* 24 (1): 83–100.

Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185 (4157): 1124–31.

Tversky, A., and D. Koehler. 1994. Support theory: A nonextensional representation of subjective probability. *Psychological Review* 101 (4): 547–67.

Walley, P. 1991. *Statistical reasoning with imprecise probabilities.* London: Chapman Hall.

Walley, P. 1996. Inferences from multinomial data: Learning about a bag of marbles (with discussion). *Journal of the Royal Statistical Society, Series B* 58 (1): 3–57.