Deference and Infinite Frames

Anonymous 2025-02-07

This paper concerns three recent results concerning probabilistic defer- ence. The results show interesting things about how various kinds of defer- ence work on finite frames, but in each case the results do not naturally gen- eralise to infinite frames. The non-generalisation raises interesting philo- sophical questions about the epistemological significance of the results, but those questions are set aside here. The priority in this paper is simply show- ing that the results fail when we allow frames to be infinite.

Recently there has been some philosophical interest in the epistemology of *deference*. Some of the important questions about deference concern pooling.[1](#_bookmark0) When we regard two other people as experts, but they disagree, how should we come up with a view that balances the two views? Some of the questions concern who is an expert. We’ve known since at least the work of David Blackwell ([1953](#_bookmark14)) that given negative and positive intro- spection for evidence, it is never a bad idea to regard one’s more informed self as an ex- pert. In general, however, negative introspection for evidence isn’t very plausible, and it would be good to know how much it can be weakened while retaining something like Blackwell’s results.

This note concerns three recent (or at least recently published) results relevant to these questions. All of the results hold for arbitrarily large finite frames. Somewhat surpris- ingly, none of them hold in general for infinite frames. I don’t have any theory about why this topic should yield so many results that have the interesting characteristic of hold- ing on all finite frames, but not all infinite frames. I also don’t have sharp bounds on

1See Elkin and Pettigrew ([2025](#_bookmark19)) for an up to date account of issues about pooling.

just when these results do and don’t hold. The examples below differ both in the cardi- nailty of the frames used (one is countable, two are uncountable), and the continuity of the value functions defined on them (one makes essential use of a discontinuous value function). So there are plenty of open questions here, although I hope there’s still some interest in these results.

# 1 Dual Deference

If A and C are probability functions, the strongest kind of deference (with respect to some proposition *p*) is when C takes A’s probability in *p* to settle what the correct prob- ability is. More formally, it is that ∀*a*: C(*p* | A(*p*) = *a*) = *a*. Our first question is when C can defer in this strong sense to two different functions A and B.

There are two cases when this can happen quite easily. The first is when C is certain that A and B will agree, i.e., C(A = B) = 1. The second is when C takes one or other of the functions to be superior, i.e., when they disagree to always go with what one particular function says. So if C takes A to be superior, then ∀*a*, *b*: C(*p* | A(*p*) = *a* ∧ B(*p*) = *b*) = *a*. But is there a third option? Can C think that A and B are both worthy of total deference, that they might disagree, and when they do the right thing to do is to land somewhere between their two credences?

Dmitri Gallow ([2018](#_bookmark20)) proved one important negative result here. He showed that there is no triple of probability functions C, A, B satisfying the following constraints.

1. ∀*a*: C(*p* | A(*p*) = *a*) = *a*;
2. ∀*b*: C(*p* | B(*p*) = *b*) = *b*;
3. C(A = B) < 1;
4. For some λ ∈ (0,1), ∀*a*,*b*: C(*p* | A(*p*) = *a* ∧ B(*p*) = *b*) = λ*a* + (1-λ)*b*.

That is, C can’t defer to both A and B individually, think that A and B might disagree, and in the event they do disagree, plan to take a fixed linear mixture of A’s probability and B’s probability as the probability of *p*. This result, unlike most I’ll discuss in this paper, does not make any finiteness assumptions, but it does make a strong assumption about mixing, namely that the mixture will be linear.

Snow Zhang ([Forthcoming](#_bookmark26)) recently proved a result that mostly generalises Gallow’s

result, though it does weaken it in one crucial respect.[2](#_bookmark2) She shows that it is impossible for A, B and C to satisfy the following five constraints.

1. ∀*a*: C(*p* | A(*p*) = *a*) = *a*;
2. ∀*b*: C(*p* | B(*p*) = *b*) = *b*;
3. C(A = B) < 1;
4. For any *a*,*b*: C(*p* | A(*p*) = *a* ∧ B(*p*) = *b*) is strictly between *a* and *b*.
5. For some finite set of values S, C(A(*p*) ∈ S ∧ B(*p*) ∈ S) = 1.

This section shows that the last constraint is essential; it is possible to satisfy the first four constraints without it. I’ll show this by constructing a model where the first four constraints are satisfied. In this model there will uncountably many values that A(*p*) and B(*p*) could take.

Let X, Y and Z be normal distributions with mean 0 and variance 1. In symbols, each of them is 𝒩(0,1). So the sum of any two of them has distribution 𝒩(0,2), and the sum of all three has distribution 𝒩(0,3). Let *p* be the proposition that this sum, X + Y + Z, is positive. Let C be a probability function that incorporates all these facts, but has no other direct information about X, Y, and Z. So C(*p*) = ½, since in all respects C’s opinions are symmetric around 0.

C knows some things about A and B. Both of them know everything C knows about X, Y, Z, and each are logically and mathematically omniscient, and know precisely what evidence they have.[3](#_bookmark3) One of them knows the value of X, and one of them knows the value of X + Y. A fair coin was flipped. If it landed heads, then A knows X and B knows X + Y; if it landed tails, it was the other way around. C knows about this arrangement, but doesn’t know how the coin landed. Let H be the proposition that it landed heads.

Since both A and B know everything C knows plus something more, and satisfy pos- itive and negative introspection, C should defer to them. If C knew which knew X + Y and which only knew X, they would defer to the one who knew X + Y. They don’t know this, but conditional on knowing the values of A(*p*) and B(*p*), they can go close to figuring it out.

2Zhang’s result is reported in Pettigrew and Weisberg ([2024](#_bookmark23)). The full result Zhang proves is considerably more general than this one, but the version I’ll discuss makes it easier to see the contrast between the finite and infinite case.

3That is, each of them satisfy positive and negative introspection for evidence. The next two sections will

drop the assumption that more informed functions satisfy negative introspection.

Assume for now that the coin landed heads, so H is true. We’ll work out the joint den- sity function for A and B. Then we can work out the same density function conditional on ¬H, and from those two facts work out the posterior probability of H. Call this value

*h*. Conditional on A(*p*) = *a*, and B(*p*) = *b*, C’s probability for p should be (1-*h*)*a* + *hb*. That’s because conditional on A(*p*) = *a*, B(*p*) = *b* and H, C’s probability for *p* should be *b*, while conditional on A(*p*) = *a*, B(*p*) = *b* and ¬H, C’s probability for *p* should be *a*. The short version of what follows is that since *h* is a function of *a* and *b* and is always in (0,1), it follows that C obeys constraint 4.

Given H, we can work out the value of X from A(*p*) = *a*. In what follows, Φ(*x*) is the cumulative distribution for the standard normal distribution, i.e., for 𝒩(0,1), and Φ-1 is its inverse. If X = *x*, then *p* is true iff Y + Z > -*x*. Since Y + Z is a normal distribution with

mean 0 and variance 2, i.e., standard deviation √2, the probability of this is Φ( √𝑥 ). So

*x* = √2Φ-1(*a*).  2

Given H, that X = √2Φ-1(*a*), and B(*p*), we can work out what Y must be as well. If

B(*p*) = *b*, that means that the probability that Z > -(X + Y) is *b*. Since Z just is a standard normal distribution, that means that X + Y is Φ-1(*b*), and hence Y is Φ-1(*b*) - √2Φ-1(*a*).

Now we can work out the joint density function for *a* and *b* conditional on H. Given H, A(*p*) = *a* and B(*p*) = *b* just when X = √2Φ-1(*a*) and Y = Φ-1(*b*) - √2Φ-1(*a*). And if

we write 𝜙(*x*) for the density function for the standard normal distribution[4](#_bookmark4), the joint

distribution for A(*p*) = *a* ∧ B(*p*) = *b* given H has density

√ √

−1 −1 −1

𝜙( 2Φ (𝑎))𝜙(Φ (𝑏) − 2Φ (𝑎))

By a parallel calculation, the joint density function for for A(*p*) = *a* ∧ B(*p*) = *b* given ¬H has density

√ √

−1 −1 −1

𝜙( 2Φ (𝑏))𝜙(Φ (𝑎) − 2Φ (𝑏))

So given that A(*p*) = *a* ∧ B(*p*) = *b*, the probability of H is

𝜙(√2Φ−1(𝑎))𝜙(Φ−1(𝑏) − √2Φ−1(𝑎))

𝜙(√2Φ−1(𝑎))𝜙(Φ−1(𝑏) − √2Φ−1(𝑎)) + 𝜙(√2Φ−1(𝑏))𝜙(Φ−1(𝑎) − √2Φ−1(𝑏))

4i.e., 𝜙(𝑥) = 𝑒√ 2 .

− 𝑥2

2𝜋

If we call that value λ, it follows that C(*p* | A(*p*) = *a* ∧ B(*p*) = *b*) = λ*b* + (1-λ)*a*, and since λ ∈ (0,1), this means that C satisfies constraint 4. This is consistent with Gallow’s result because λ is not a constant, it is a function of *a* and *b*. And it is consistent with Zhang’s result because each of A(*p*) and B(*p*) can take infinitely many, in fact uncountably many, values. If one tries to make a similar construction to this one with only finitely many possible values for the probabilities, there will be some value which only the more in- formed probability can take, and in that case C’s posterior probability will be equal to the probability of the more informed expert.

To understand the relationship between *a*, *b*, and C’s posterior probability, it helps to visualise one part of it. Figure [1](#_bookmark5) shows what value this posterior takes for different values of *b* holding fixed *a* = 0.75.

1.00

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

0.75

Posterior value of C(*p*)

0.50

0.25

0.00

0.00 0.25 0.50 0.75 1.00

B(*p*)

Figure 1: The posterior probability of C(*p*) given A(*p*) = 0.75.

The distribution loosely follows what Levinstein ([2015](#_bookmark22)) calls Thrasymachus’s Princi- ple. The more opinionated of the two experts gets much stronger weight. You can see this in part by seeing how close the above graph gets to *x* = *y* at either extreme. But it’s perhaps more vivid if we plot the posterior probability that the coin landed Tails against

the different values of B(*p*), as in Figure [2](#_bookmark6).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

0.6

Posterior probability of Tails

0.4

0.2

0.0

0.00 0.25 0.50 0.75 1.00

B(*p*)

Figure 2: The posterior probability of the coin landing Tails given A(*p*) = 0.75.

When B(*p*) is between 0.25 and 0.75, i.e., when it is closer to 0.5 than A(*p*) is, C is confident that the coin landed Tails, and that A is more informed and hence more worthy of deference. When B(*p*) takes a more extreme value, then C is confident that the coin landed Heads, and hence that B is more worthy of deference. In general, this model backs up Levinstein’s intuition that more opinionated sources are probably better informed, and hence more worthy of deference.

# 2 Evidence and Nesting

The previous section assumed that C strongly deferred to A and B. I now turn to the question of when C should do that. A natural thought, one I relied on in that discussion, was that C should defer when they regard A and B as better informed than they are. This can be motivated with a famous result from David Blackwell ([1953](#_bookmark14)). Let E1 and E2 be

functions from W to subsets of W. Intuitively, these are *experiments*; The Experimenter will perform E*i* and learn they are in E*i*(*w*), where *w* is the world they are in. Blackwell assumes that the range of each E*i* is a partition of W; The Experimenter always learns what cell of the partition they are in.

The short version of the big result is that E1 is guaranteed to be more valuable than E2 iff E1 is more informative than E2. All of that needs clarifying though.

Say E1 is a *refinement* of E2 iff for all *w*, E1(*w*) ⊆ E2(*w*). Formally, this is how I’ll capture the intuitive notion of being more informative.[5](#_bookmark8)

Let O be a finite set of options: {O1, …, O*n*}. Each O*i* is a function from W to reals. Intuitively, they are bets, and the number is the return on each bet. I’ll follow standard terminology in philosophy and say that a function from worlds to reals is a *random vari- able*. Given a random variable X (defined on W) and a probability function Pr, we can define the expectation Exp(W, Pr) as ΣPr(*w*)X(*w*), where the sum is across members of W.[6](#_bookmark9)

Say a strategy S is a function from E and W to O such that if E(*x*) = E(*y*), then S(*x*) = S(*y*). That is, strategies are not more fine-grained than evidence. Intuitively, a strategy is something that The Experimenter can implement given their evidence, so it can’t require them to make more discriminations than their evidence does. For each S, we can define a random variable SR (read this as the return of S), such that Sr(*w*) = S(*w*)(*w*). In words, the return of S at *w* is the value at *w* of the option S selects at *w*.

Finally, say that a strategy is **recommended**[7](#_bookmark10) by Pr (relative to E, O and W) just in case for all *w* in W, and alternative options Oa in O, Exp(S(*w*), Pr(• | E(*w*))) ⩾Exp(Oa, Pr(• | E(*w*))).

In words, the option selected by S at *w* has maximal expected utility out of the options in O, relative to the result of updating Pr on the evidence available at *w*. Given these

5Note that comparisons here always include equality. An experiment is a refinement of itself, is more informative than itself, and is more valuable than itself. This can lead to confusion, but it’s the standard terminology, and the alternative is much more wordy.

6I’m defining random variable and expected value the way they are usually defined in philosophy. In

some fields it is more common to define random variables as functions from a probability space to reals, where a probability space has W and Pr as constituents. Then we can define expectation as a one- place function that simply takes a random variable as input. I think the philosophers’ way of speaking is more useful, and in any case I’m a philosopher so it’s more natural to me. But note there is a potential terminological confusion here.

7This term is taken from Dorst et al. ([2021](#_bookmark18))

notions, we can state two important results Blackwell proves.

First, for any O, W and Pr, if E1 is a refinement of E2, S1 is recommended by E1 and S2 is recommended by E2, then Exp(S1, Pr) ⩾ Exp(S2, Pr). No matter what practical problem The Experimenter is facing, and no matter what their priors are, they are better off adopting a strategy recommended by the more informative experiment.

Second, for any W, if E1 is not a refinement of E2, then for some O and Pr, there exists an S1 is recommended by E1 and S2 is recommended by E2 such that Exp(S1, Pr) < Exp(S2, Pr). That is, if E1 is not more informative than E2, then for some practical problem, it is better in expectation to perform E2 and carry out some strategy recommended by it.

Blackwell isn’t the first to connect the value of experiments to the relative value of strategies and options in this way; for some history of this idea see Das ([2023](#_bookmark16)) and Cam ([1996](#_bookmark15)). But he works out the consequences of it in much more detail than anyone had before him. (The two results I’ve listed here don’t go close to exhausting what he proved, but they are enough for our purposes.) Philosophers have primarily focussed on the first of these two results. And they have focussed largely on the special case when E2(*w*) = W; i.e., when ‘performing’ experiment E2 means getting no information at all.[8](#_bookmark11) In this sec- tion we’ll also ignore the second result, but we will pay attention to the case where E2 isn’t universal.

John Geanakoplos ([[1989] 2021](#_bookmark21)) proved an interesting generalisation of this first re- sult. As noted earlier, Blackwell’s theorems presuppose that each experiment is parti- tional. Formally, E is partitional iff for all *w*, *w* ∈ E(*w*), and for all *w*, *v*, if *v* ∈ E(*w*), then E(*w*) = E(*v*). Geanakoplos shows something interesting about experiments that are re- flexive, transitive, and nested. These are defined as follow (with leading quantifiers over worlds left implicit)

**Reflexive** *w* ∈ E(*w*).

**Transitive** If *v* ∈ E(*w*), then E(*w*) ⊆ E(*v*).

**Nested** Either E(*w*) ∩ E(*v*) = ∅, or E(*w*) ⊆ E(*v*), or E(*v*) ⊆ E(*w*).

He shows that both of Blackwell’s results continue to hold when E1 is reflexive, tran- sitive, and nested, as long as E2 is partitional.

8Further, they haven’t always credited Blackwell (or the earlier results from Peirce and Ramsey that Das discusses) when they do discuss the results. I’m grateful to John Quiggin for pointing out to us the importance of Blackwell’s results in this context.

This result has had some influence on recent philosophical work. It suggests the fol- lowing kind of argument.

1. Performing experiments is valuable.
2. Performing experiments is valuable iff experiments are nested.
3. Therefore, experiments are nested.

That’s fairly crude as stated, but it’s possible to develop it into a more sophisticated argument that has implications for what the correct epistemic logic should be. You can (more sophisticated) versions of this argument in Spencer ([2018](#_bookmark24)) and Dorst ([2019](#_bookmark17)), and criticisms of these arguments in Williamson ([2019](#_bookmark25)) and Das ([2023](#_bookmark16)).

The aim of this section is to show that two of the assumptions that Geanakoplos uses in proving these results are essential. First, E2 has to be partitional; it is not sufficient that it is reflexive, transitive, and nested. Second, it cannot be that both W and O are infinite. Both results arguably help the Williamson-Das side of the debate mentioned in the previous paragraph, but I won’t go into that in more detail here. The aim is just to show the formal limits of Geanakoplos’s result.

Here is the model that shows that more refined experiments do not necessarily have higher expected value if both experiments are reflexive, transitive, and nested.

* + W = {w1, w2, w3}
  + Pr(w1) = Pr(w2) = Pr(w3) = ⅓
  + O = {O1, O2}
  + O1(w1) = O1(w2) = O1(w3) = 0
  + O2(w1) = 3
  + O2(w2) = 9
  + O2(w3) = -6
  + E1(w1) = {w1, w3}
  + E1(w2) = {w2}
  + E1(w3) = {w3}
  + E2(w1) = {w1, w2, w3}
  + E2(w2) = {w2, w3}
  + E2(w3) = {w3}

Given no information, the optimal strategy is to always take the bet, i.e., choose O2 over the fixed return of 0 that is O1. This has an expected return of 2. Given E1, the

only recommended strategy is to choose O1 at w1 and w3, and O2 at w2, for an expected return of 2. But given the less informative E2, the recommended strategy is to choose O2 at w1 and w2, and O1 at w3, for an expected return of 4. In this case, performing the less informative experiment has higher expected returns. (Though to be clear, both experiments have positive expected returns, relative to not doing anything.)

Next I’ll show the result does not hold when W and O are infinite. It pays to be careful here because it’s easy to have a case where S1 and S2 are undefined. And it’s not interesting that Exp(S1, Pr) ⩾ Exp(S2, Pr) might sometimes fail to be true simply because one or other term in it isn’t defined. So I’ll restrict attention to cases where utilities are bounded, for any E and any *w* there is an optimal strategy given E(*w*), and the expectation of any recommended strategy is defined.

Let W be the reals in (0,1). E1(*x*) = [*x*,1), and E2(*x*) = (0,1). For any *x* in [0,1], let O*x*(*y*) be -1 if *x* ≥ *y*, and *x* if *x* > *y*, and O be the set of all these O*x*. And Pr is the flat distribution over (0,1); the only fact I’ll need is that whenever 0 < *x* < *y* < 1, the probability that the actual world is in (*x*, *y*) is *y* - *x*.

The only strategy recommended by E2 is to always choose O0, which has a guaranteed return of 0. The strategy recommended by E1 is to choose O*x* upon learning that the true value is in (*x*, 1). That has an expected return of *x*, which is higher than all the alternatives. But following that strategy all the time has a guaranteed return of -1, which is worse than the strategy recommended by E2. And that’s true even though E2 is partitional, and E1 is reflexive, transitive, and nested.

Now there is something odd about this example - each of the O*x* is discontinuous. In each case, the payouts jump from -1 to *x* at a particular point. This suggests a further open question: If every member of O is continuous, are more refined experiments valuable in expectation? It’s also, as far as I know, open whether Geanakoplos’s results hold in countable frames. But what this model shows is that his result does not generalise to uncountable frames with discontinuous payouts.

# 3 Trust and Value

The role of experiments in the frames discussed in Section [2](#_bookmark7) is somewhat curious. They are in one respect central, the theorems are all restricted to whether frames are partitional, nested, etc, but they are in another respect ephemeral. Ultimately what matters is not the

experiment, but the probability that The Experimenter has after performing an experi- ment. This latter way of thinking is a helpful way to understand a striking recent result by Dorst et al. ([2021](#_bookmark18)).

Say a probability frame is an ordered pair ⟨W, P⟩ such that W is a set (intuitively, of worlds), and P is a function from W to probability functions defined on w. One way to generate such a pair is to have some experiment E and prior probability Pr, each defined on W, and have P(*w*) be Pr(• | E(*w*)). But you can just cut out the E and Pr, and focus simply on W and P. As Dorst et al. ([2021](#_bookmark18)) show, this turns out to be mathematically a very helpful move. It lets you see a lot more interesting features of these frames.[9](#_bookmark12) They call frames where P is generated from E and Pr in this way *prior frames*, and those will be the focus of discussion here, but it is interesting to see them as a special case of a more general class.

One thing they prove by looking at the more general class is that when W is finite, the following two claims are equivalent. (I’ll state the claims formally, then explain my notation.)

**Total Trust** E(X | {*w*: Exp(X, P(*w*)) ⩾ *t*}, π) ⩾ *t*

**Value** If O is a set of options, *s* is a recommended strategy for O, and *o* is a member of O, then Exp(*s*, π) ⩾ E(*o*, π).

There is a bit there to unpack. I’ll follow them in using π for a probability function that is outside the frame. I’ll sometimes call it Novice’s probability function, as opposed to P(*w*) which is The Experimenter’s probability function at *w*.[10](#_bookmark13) I’m generalising the notion of expectation a bit to allow for conditional expectations; Exp(X | p, Pr) is the expectation of X according to Pr(• | p). So here’s what Total Trust says. Take any random variable X. Update π on the proposition that consists of all and only worlds where The Experimenter at that world has an expected value for X at least equal to *t*. After that update, Novice also expects X’s value to be at least *t*.

I discussed recommended strategies in Section [2](#_bookmark7), so there is less to say about Value. What it says is that Novice does not expect any member of O to do better than any rec- ommended strategy.

9At least in the case where W is finite; I’ll be getting to the issues when it is not.

10When the frame is a prior frame, it is natural to focus on the case where π = Pr, but again I’m not just looking at prior frames.

One way Dorst et al put the equivalence between Total Trust and Value (on finite frames) is that π Totally Trusts a frame ⟨W, P⟩ iff it Values that frame. What I’ll show is that this equivalence breaks down without the finiteness assumptions. Indeed, it breaks down even when W is countably infinite.

Start with a frame I’ll call **Coin**. A fair coin will be flipped repeatedly until it lands Tails. Let F be a random variable such that F = *x* iff the coin is flipped *x* times. (If the coin never lands Tails, I’ll stipulate that F = 1. Since this has probability 0, it doesn’t make a difference to the probabilities, but it will make a difference to the possible choices post- update.) Novice knows these facts about F, so π(F = *x*) = 2-*x* . If F = *x*, then The Experi- menter learns F ⩾ *x* and nothing else, and updates on that. That is, P(F = *x*) = π(•|F ⩾ x). For any positive integer *i*, let Oi be the random variable that takes value 0 at F = *j* when *j* ⩽ *i*, and value 2*i* at F = *j* when *j* > *i*. Let O be the set of each O*i*. The strategy *s* such that *s*(F = *i*) = O*i* is recommended, as can be easily checked. But E(*s*) = 0, while for any *o* ∈ O, E(*o*, π) = ½. So Value fails on **Coin**.

On the other hand, π does Totally Trust **Coin**. For any random variable X and thresh- old *t*, say an integer *k* is a cut-off if either Exp(X | F ⩾ *k*, π) ⩾ *t* and Exp(X | F ⩾ *k* + 1, π) < *t*, or Exp(X | F ⩾ *k*, π) < *t* and Exp(X | F ⩾ *k* + 1, π) ⩾ *t*. Let c*i* be the *i*’th cutoff. Partition the integers into the regions between cutoffs. More precisely, do the following. If 1 is not a cutoff, the first cell of the partition is {1, …, c1-1}; otherwise the first cell is just {1}. If there is a last cutoff *c*, the last cell is {c, c+1, …}. Otherwise, each cell is {c*i*, …, c*i*+1-1}. Say a cell is *positive* if for every *k* in it, Exp(X | F ⩾ *k*, π) ⩾ *t*, and negative otherwise. (By the construction of the cells, Exp(X | F ⩾ *k*, π) ⩾ *t* is true for either all or none of the members.)

Let {c*i*, …, c*i*+1-1} be an arbitrary positive cell. By construction, Exp(X | F ⩾ c*i*, π) ⩾ *t*, and Exp(X | F ⩾ c*i*+1, π) < *t*. Since for some λ ∈ (0,1), Exp(X | F ⩾ c*i*, π) = λExp(X | F ∈

{c*i*, …, c*i*+1-1}, π) + (1-λ)Exp(X | F ⩾ c*i*+1, π), it follows that Exp(X | F ∈ {c*i*, …, c*i*+1- 1}, π) ⩾ *t*. Since this was an arbitrary positive cell, it follows that Exp(X | F ∈ I, π) ⩾ *t* for any positive cell I. Since Exp(X | {*w*: Exp(X, P(*w*)) ⩾ *t*}, π) is a weighted average of the values of Exp(X | F ∈ I, π) where I is one or other of the positive cells, it follows that E(X | {*w*: Exp(X, P(*w*)) ⩾ *t*}, π) ⩾ *t*, as required.

So π Totally Trusts **Coin**, but doesn’t Value it. So the equivalence between Total Trust and Value fails here. But you might very reasonably object on two scores. First, the value function used to generate the counterexample was unbounded, and we know that un-

bounded value functions lead to all sorts of paradoxes. Second, I didn’t just make W infinite, I made O infinite as well, so this isn’t a minimal generalisation of the original claim. It turns out that if we put both these constraints on, then the equivalence fails in the other direction: It is possible to get a frame that π Values, but does not Totally Trust.

Call the following frame **Bentham**. Again, a coin will be flipped until it lands Tails. If it ever lands Tails, F is the number of flips. If it never lands Tails, which has probability 0, then F = ∞. Again, Novice knows these facts, and so far the case is just like **Coin**. But in this case, if F = *x*, The Experimenter learns that F ⩽ *x* and nothing else, and The Exper- imenter updates on that. So if F = ∞, The Experimenter learns nothing, but otherwise they can rule out all but finitely many possibilities. More precisely, P(F = *x*) = π(•|F ⩽ x). The Novice probability does not Totally Trust this frame. Let Y be a ran- dom variable such that Y(F = ∞) = 0, and for all finite *n*, Y(F = *n*) = 1 - 2-*n*. E(Y | {*w*: Exp(Y, P(*w*)) ⩾ ⅔}, π) = 0 < ⅔. The only world *w* where Exp(Y, P(*w*)) ⩾ ⅔ is

F = ∞, and at F = ∞, Y = 0.

On the other hand, π does Value this frame. To see this, for any set of options O, recommended strategy S, random variable X (all defined on W), and integer *n*, let W*n* be the set {F = 1, …, F = *n*}, O*n*, P*n*, S*n* and X*n* be the restrictions of O, P and S to worlds in W*n*. From the way P is constructed, i.e., by conditionalising on the set of worlds where F is no greater than it actually is, it follows that if S is recommended on ⟨W, P⟩, then S*n* is recommended on ⟨W*n*, P*n*⟩. Since ⟨W*n*, P*n*⟩ is a finite prior frame where E is reflexive, transitive and nested, and Pr = π, it follows by the result of Geanakoplos described in Section [2](#_bookmark7), that the expected return of S*n* is greater than the expected return of any option in O*n*. For any random variable X, Exp(X, π) is the limit as *n* tends to ∞ of Exp(X*n*, π); this is because as *n* grows this covers all words in W except F = ∞, which has probability

0. If the expected return of S is the limit *n* tends to ∞ of the expected return of S*n*, and the expected return of an option in O is the limit as *n* tends to ∞ of its counterpart in O*n*, and S*n* is better (in expectation) than every option in O*n*, it follows that S is better (in expectation) than every option in O. So Value is satisfied, as required.

# 4 Conclusion

It’s striking that we get such different behaviours between finite and infinite frames when it comes to these three somewhat distinct issues about deference and updating. The main

point of this note is to point out these differences.

But there is a broader philosophical question. One might think that since humans are finite, results that hold on all finite frames should be used when thinking about humans. I think this is a bit quick. All the models I’ve been using here, both the finite and the in- finite ones, really are *models*. Even the finite ones assume that the people being modeled have superhuman (if not literally infinite) computational capacities. They are all idealisa- tions. The question is, are they good or bad idealisations? Here the issues about finitude get complicated. It might be that an infinite model is a better idealisation, a better ap- proximation to a human than a finite one.

Imagine someone saying that since humans are finite, and circles are infinitely curved, we should never model humans as thinking about circles. Rather, we should think that the human is thinking of a regular polygon with arbitrarily many sides. This is a little absurd. A model where the human is thinking of a circle is simpler than a model where the human is thinking, in full precision, about a chiliagon. By the same reasoning, it might be better to model a scientist as taking some variable to be normally distributed over an interval, than to take them to have in their head a particular finite approximation to the normal distribution, which they perfectly update. For that reason, the model in Section [1](#_bookmark1) might be a decent model of deference.

The models in Section [2](#_bookmark7) are, admittedly, weirder. There is less use, even in idealisations, for infinite models with discontinuous payouts, or for models with unbounded utility functions. These are known to lead to weird results. Here I think there is a stronger claim that the models I’ve presented are not useful models, not because they are infinite, but because they are discontinuous and unbounded.

But there is obviously much more to say about these questions about usefulness. Hopefully it is helpful to simply point out how differently the finite and infinite cases behave.

**References**

Blackwell, David. 1951. “Comparison of Experiments.” *Proceedings of the Berkeley Sym-* *posium on Mathematical Statistics and Probability* 2 (1): 93–102.

———. 1953. “Equivalent Comparisons of Experiments.” *The Annals of Mathematical* *Statistics* 24 (2): 265–72.

Cam, L. Le. 1996. “Comparison of Experiments: A Short Review.” In *Statistics, Proba-*

*bility and Game Theory: Papers in Honor of David Blackwell*, edited by T. S. Fergu- son, L. S. Shapley, and J. B. MacQueen, 127–38. Hayward, CA: Institute of Math- ematical Statistics. <https://doi.org/10.1214/lnms/1215453569>.

Das, Nilanjan. 2023. “The Value of Biased Information.” *British Journal for the Philos-* *ophy of Science* 74 (1): 25–55. <https://doi.org/10.1093/bjps/axaa003>.

Dorst, Kevin. 2019. “Evidence: A Guide for the Uncertain.” *Philosophy and Phe-* *nomenological Research* 100 (3): 586–632. <https://doi.org/10.1111/phpr.12561>.

Dorst, Kevin, Benjamin A. Levinstein, Bernhard Salow, Brooke E. Husic, and Branden Fitelson. 2021. “Deference Done Better.” *Philosophical Perspectives* 35 (1): 99–150. <https://doi.org/10.1111/phpe.12156>.

Elkin, Lee, and Richard Pettigrew. 2025. *Opinion Pooling*. Cambridge: Cambridge Uni- versity Press. <https://doi.org/10.1017/9781009315203>.

Gallow, J. Dmitri. 2018. “No One Can Serve Two Epistemic Masters.” *Philosophical* *Studies* 175 (10): 2389–98. <https://doi.org/10.1007/s11098-017-0964-8>.

Geanakoplos, John. (1989) 2021. “Game Theory Without Partitions, and Applications to Speculation and Consensus.” *The B.E. Journal of Theoretical Economics* 21 (2): 361–94. https://doi.org/<https://doi.org/10.1515/bejte-2019-0010>.

Levinstein, Benjamin Anders. 2015. “With All Due Respect: The Macro-Epistemology of Disagreement.” *Philosophers’ Imprint* 15 (13): 1–20.

Pettigrew, Richard, and Jonathan Weisberg. 2024. “Updating on the Evidence of Oth- ers.” *Philosophical Studies* 181 (10): 2539–62. [https://doi.org/10.1007/s11098-](https://doi.org/10.1007/s11098-024-02173-z) [024-02173-z](https://doi.org/10.1007/s11098-024-02173-z).

Spencer, Jack. 2018. “No Crystal Balls.” *Noûs* 54 (1): 105–25. [https://doi.org/10.](https://doi.org/10.1111/nous.12252)

[1111/nous.12252](https://doi.org/10.1111/nous.12252).

Williamson, Timothy. 2019. “Evidence of Evidence in Epistemic Logic.” In *Higher-Order Evidence: New Essays*, edited by Mattias Skipper and Asb- jørn Steglich-Petersen, 265–97. Oxford: Oxford University Press. [https:](https://doi.org/10.1093/oso/9780198829775.003.0013)

[//doi.org/10.1093/oso/9780198829775.003.0013](https://doi.org/10.1093/oso/9780198829775.003.0013).

Zhang, Snow. Forthcoming. “Coherent Combinations of Multiple Experts’ Opinions: Another Impossibility Result.” *Theory and Decision*, Forthcoming.