Darren Bradley

Four Problems about Self-Locating Evidence

May 2007

Introduction

The strategy of this paper is unify and conquer. I will show that four problems that appear to be very different have the same structure. I give a unified treatment of the Doomsday Argument, Sleeping Beauty, the Fine-tuning Argument and confirmation in the Everett interpretation of quantum mechanics. All these cases involve self-locating evidence. However, the confusing feature of all these cases is not self-location, but observation selection effects. I explain how observation selection effects operate, why they affect the four problem cases, and how they can be incorporated into confirmation theory. I will defend the Doomsday Argument, the halfer position in Sleeping Beauty, the Fine-tuning Argument and the applicability of confirmation theory to the Everett interpretation of quantum mechanics.

Uncertain Aces

Let's start with the Bayesian framework we will be using. Assume that all your beliefs have a degree of certainty between 0 and 1. Such beliefs can be modelled with a probability function. We will use the following probability-raising account of confirmation:

E confirms H if and only if $P(H|E) > P(H)$

P(H|E) stands for the conditional probability of H given E. Conditional probability is defined as follows:

$$P(H \mid E) = \frac{P(H \& E)}{P(E)}$$

A useful theorem is the following:

$$P(H|E) > P(H) \text{ if and only if } P(E|H) > P(E|-H)$$

$$\text{if and only if } P(E|H) > P(E)$$

Consider the following probability problem:

Uncertain Aces

Alice is dealt one or two cards, determined by the flip of a fair coin. One card is dealt if Heads lands; two if Tails lands. If two cards are dealt, one is an Ace and one is a King. If one card is dealt, a further coin is flipped to decide if an Ace or King is dealt.

	Card 1	**Card 2**
Heads	Ace or King	-
Tails	Ace	King

Assume your prior probability of Heads is 50%.

$P(H) = 1/2$

P(An Ace is dealt | Heads) = 1/2

P(An Ace is dealt | Tails) = 1

Alice now tells you about one of her cards:

E = Alice says ‘I have an Ace’.

Does E confirm Heads, Tails or neither?

E confirms H if and only if $P(H|E) > P(H)$. Recall that $P(H) = 1/2$.

$$P(H \mid E) = \frac{P(H \,\&\, E)}{P(E)}$$

We need values for P(E) and P(H&E) to find out if E confirms H. But we have not yet been given enough information to determine of E confirms H.

Ontic and Epistemic Processes

To see the problem, ask the following question: by what processdid I find out that an Ace had been dealt? If you have found a piece of evidence, there must be some process, some mechanism, by which the evidence was found. The mechanism in this case is that Alice tells you she has been dealt an Ace. But this leaves the situation under-specified, for we haven't been told the process he used to tell us about the Ace. Why has she told us about the Ace? Was she asked, or was it volunteered? If it was volunteered, did she particularly want to tell us about the Ace? Or would she rather have told us about a King? The answer to these questions is important in assessing the significance of what we have been told. Let's consider two simple decision procedures Alice might have used.

> Random: Alice picks a card at random from her hand and tells you what it is. If she is dealt one card she just tells you what it is (she has no choice). If she has two cards, she flips a coin to determine which one she will tell you about.

> Persistence: Alice looks for an Ace. If she finds it, she tells you she has an Ace. (If she doesn't have an Ace, she picks a card at 'random' (as above) from her hand.)

The process will have an important impact on the inferences we can draw[i]. Let's go through both cases.

Persistent Process

Suppose the process used is *persistence*. Alice looked for an Ace, found one, then told you about it. Does this disconfirm Heads? Yes. Alice wants to tell you about an Ace. If Tails lands, she can definitely do so, as she is certain to have an Ace. But if Heads lands, Alice may have only been dealt a King. Then she cannot tell you about an Ace. So if

Alice does successfully announce that she has an Ace, Tails is confirmed. (This is because the probability of the evidence given Tails is greater than the probability of the evidence given Heads.) We can see this result mathematically and diagrammatically. Recall that H = Heads, and E = Alice says ‘I have an Ace’.

$$P(H \mid E) = \frac{P(H \,\&\, E)}{P(E)}$$

$$= \frac{1/4}{3/4} = \frac{1}{3}$$

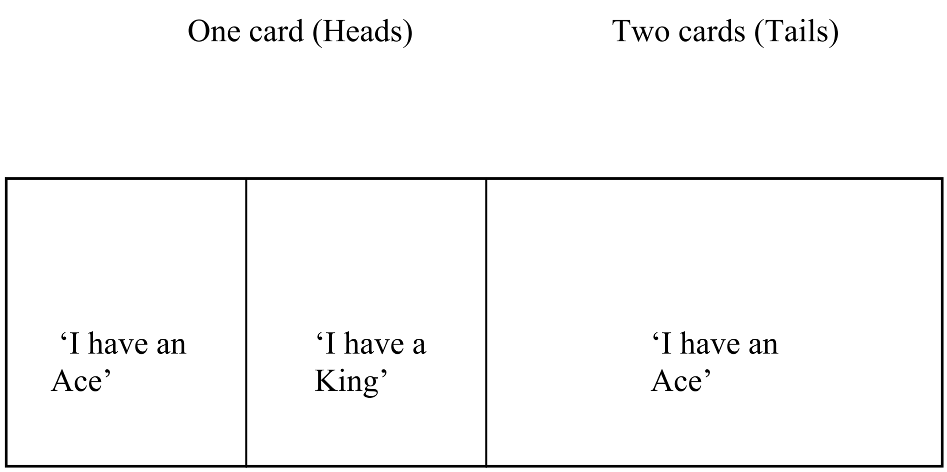

Figure1: Persistent Process

Figure 1 represents the evidence you might get on either hypothesis. If Tails, you are guaranteed to be told ‘I have an Ace’. If Heads, you will be told about an Ace with probability 1/2. (Notice figure 1 does not show the cards *dealt,* but the cards found out about.) If the evidence is persistent and there are two cards, you will be told about an Ace for certain. So being told that there is an Ace confirms Tails and disconfirms Heads.

The Ace confirms Tails because of two facts. Firstly, Tails makes it more likely that an Ace will be dealt at all. Secondly, the persistent process ensures that the more cards there are, the greater the chance that an Ace is selected. This two-stage procedure is at the heart of observation selection effects. Let's look at these processes more closely.

First there is an *ontic* process, which in this case results in the Ace being dealt. Call the outcome of the ontic process o for outcome i.e., an Ace is dealt. I will use 'outcome' as a semi-technical term for the result of the ontic process. The outcome is a set of concrete objects; it may be a set of days, people, universes or, in this case, cards. I will refer to a single object in this set as 'an outcome' or 'one of the outcomes'.

Second, there is an *epistemic* process by which the observer learns that an outcome exists that has a certain property. The property may be being a particular day (being Monday), having a particular birth rank, or, in this case, being an Ace. The epistemic process can be thought of as a relation between a property and a piece of evidence. A piece of evidence is something that might be believed[ii]. If a piece of evidence is persistent with respect to a property, it means that if the property is instantiated among the outcomes, this fact will be reported in the evidence. The epistemic process can also be thought of as a function from outcomes to evidence. If property p is instantiated by one of the outcomes and the process is persistent, then the fact that property p is instantiated will be expressed in the evidence. We will sometimes talk about an outcome being persistent (or random). This means that the property the outcome has, and which we use to refer to that outcome, is persistent (or random) with respect to the evidence.

Random Process

Let's go back to the cards. Suppose the process was random. Alice just picked a card at random and told you what it was. If the coin landed Heads, there is a 1/2 probability of her telling you about an Ace or a King. (This is the same as if the process is persistent.) If the coin landed Tails, she could have told you she had an Ace or a King with equal probability. She has both the Ace and King, and she picks one by flipping a fair coin. In this case E doesn't confirm Heads.

$$P(H \mid E) = \frac{P(H \,\&\, E)}{P(E)}$$

$$= \frac{1/4}{1/2} = \frac{1}{2}$$

The change compared to having persistent evidence is the value of P(E). This was greater than 1/2 when the evidence was persistent, as Alice was biased towards telling you about an Ace. But when the process is random, being told about an Ace is as likely as being told about a King i.e. probability 1/2.

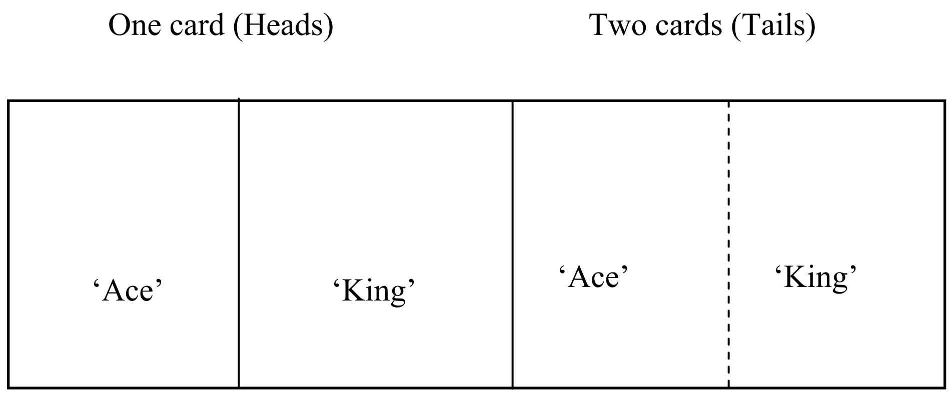

Figure 2: Random Process

The dashed line represents that Alice can choose which card to tell you about. The diagram shows that if there are two cards, there is a 50% chance you'll be told about an Ace. And if there is one card, there is a 50% chance you'll be told there is an Ace. So the evidence gives you no helpful information. Your degree of belief in Heads stays at 1/2 as it was before. There is a difference between what *happens* given Heads or Tails of course. If Tails landed, there are two actual cards, but you have only been told about one of them. If Heads landed, there is only one card, and you have been told what it was. Although the cases are different, the effect of the evidence is exactly the same: none – neither hypothesis is confirmed.

We can now state an important result. If an Ace is certain to be dealt, and the process is persistent, then finding an Ace confirms Tails (the many cards hypothesis). If the process is random, then finding an Ace does not confirm either hypothesis. This applies generally. If a particular outcome is persistent, then finding that outcome confirms the Many hypothesis. If that outcome is not persistent, then finding it does not confirm any hypothesis. This result is summarised and generalised below. Let Tails be assimilated into Many Outcomes, and Heads into Few Outcomes. Let p be the property instantiated by the outcome learnt about (being an Ace).

	p is not certain to be instantiated
Random process	(1) No confirmation
Persistent process	(2) Many Outcomes confirmed

Table 1

Persistence is a natural assumption to make, and is generally made by Bayesians without anyone realizing that it is a substantive assumption. The reason it is such a natural assumption is that it always holds if the following condition is met:

(U) For any given hypothesis, there will only be one outcome, o.

Recall that o is the outcome of the ontic process. Call this condition U for 'unique outcome'. This condition is satisfied if Alice is only dealt one card. Then there can be no funny business about which card she announces. If an Ace is dealt, we are told; and if a King is dealt, we are told. Any epistemic process will be a trivial one in which we simply find out which card has been dealt. But Tails results in two outcomes: an Ace being dealt and a King being dealt. This means there are two possible pieces of evidence – 'I have an

Ace' or 'I have a King'. Once we have more than one outcome for a given hypothesis, we have to know the process by which the evidence we have was found. Otherwise it is impossible to work out what effect it has on the probability of the hypothesis.

Note that (U) doesn't imply that the prior probability of a given outcome is 1 or 0. There may still be a non-trivial probability distribution, as there is if Heads (an Ace or a King may be dealt). (U) merely says that only one outcome will be *actual.* This condition is satisfied for Heads, despite the non-trivial probability distribution over outcomes. But the condition (U) is not satisfied for Tails, even though Tails does have a trivial probability distribution over outcomes (an Ace and a King will both be dealt with probability 1). We will see that (U) is generally not satisfied when self-locating evidence is learnt.

The moral is that we must be careful how we apply conditionalization. As written, conditionalization makes no mention of the process. It simply tells us to conditionalize on the evidence learnt. This is univocal if condition (U) holds. But if (U) doesn't hold, as in our cards case, conditionalization seems to under-specify what we should do. It doesn't tell us how to take the process into account. So what should we do?

This question is at the root of studies of observation selection effects (see Bostrom 2002 for a book-length study). But I think there is a simple solution regarding what we should do. We should conditionalize on the original evidence, E, *plus* the process by which E was found. Call this combined evidence E*. It is this more detailed piece of evidence that we should conditionalize on (see Hutchison 1999).

To re-cap, whenever we learn about a particular outcome, there is a two-stage process. There is some ontic process which results in that outcome occurring, and there is an epistemic process by which we come to learn about that outcome. Both of these processes are an essential part of any inference we can draw. E* represents the total evidence once both of these effects have been taken into account.

Certain Aces

It will be useful to consider one more card case. We assumed in *Uncertain Aces* that if Heads landed, either an Ace or a King could be dealt. Let's now alter this so that an Ace is dealt if Heads lands. All else is as before.

Certain Aces

	Card 1	**Card 2**
Heads	Ace	-
Tails	Ace	King

This dramatically changes the inferences we can draw when we learn E = Alice says 'I have an Ace'. The property the observed outcome has – being an Ace – is now certain to be instantiated.

If the outcome was persistently discovered, and an Ace was searched for, then E has a probability of 1. So it doesn't favour either hypothesis (box 4). If the outcome was randomly discovered, then Heads is confirmed because Heads entails the outcome will be found, but Tails only assigns a 50% probability to the outcome being found (box 3).

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No confirmation	(3) Few Outcomes confirmed
Persistent process	(2) Many Outcomes confirmed	(4) No confirmation

Table 2

We saw earlier that when p is not certain to be instantiated, Tails is confirmed if being an Ace is persistent (box 2) and not if it is random (box 1).

Table 2 is a map of this paper. I will now show how each of four problems of self-locating evidence fit into each of the four boxes. They all have the following structure:

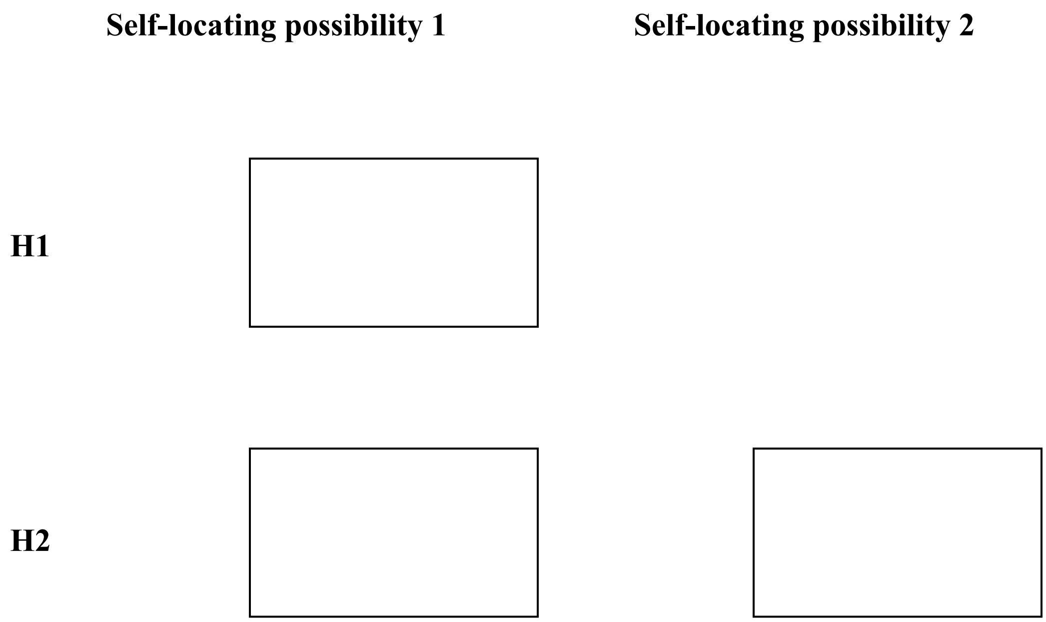

Figure 3: Self-locating possibilities

There are two possible worlds, with two possible locations in each world. The question is whether learning that some property is instantiated confirms H1 or H2. We will see that that depends on two variables – the epistemic process by which the property was discovered, and whether the property discovered was certain to be instantiated. The four problems correspond to each of the four self-locating variables – agent, time, space and branch.

The Doomsday Argument

Let's start with the problem in which the self-locating variable discovered is which agent you are; this is the Doomsday Argument[iii]. A simplified version will be all we need. We can generate the simplest possible Doomsday-style argument by imagining that there is either a total of one or a total of two people in the universe. These hypotheses are mutually exclusive and exhaustive.

H1: There is one person in the universe (Doom Soon)

H2: There are two people in the universe (Doom Late)

Assume each has a probability of 50%. Each person created is put in an isolation cubicle, so they do not know if anyone else exists. They are numbered: person 1, and if he exists, person 2. Suppose you find yourself existing in this scenario (which you know about). There are three possible states you might be in.

	Self-Locating Possibility 1	**Self-Locating Possibility 2**
H1	Person 1	
H2	Person 1	Person 2

Figure 4: The Doomsday Argument

(Note the similarity to *Certain Aces.)* Suppose you learn that you are person 1. Does this confirm H1, H2 or neither. We must ask about the two variables: by what process have I discovered the evidence, and is the property discovered certain to be instantiated?

Start with the process. The ontic process results in either one or two people existing. The epistemic process results in you observing that either person 1 or person 2 exists. You do this by *being* person 1 or person 2, and then learning your birth rank. Let's state the processes from earlier in more general terms than previously.

> Random: An outcome is picked at random from the actual outcomes. There is no bias in the selection process.

> Persistence: If an outcome with a particular property exists, the observer is told about that outcome. (If it doesn't exist, an outcome is picked at random.)

Notice these processes only differ when there is more than one outcome. In this case, the processes only differ if there is more than one person in existence (otherwise there is only one person for the epistemic process to find). So the processes only differ if there are two people. Supposing there are two people, are you more likely to discover that one of them exists rather than the other? This amounts to: supposing there are two people, are you more likely to be person 1 than person 2? The answer is no. There is no *a priori* bias towards being one of the people rather than the other. In the Doomsday Argument literature, this is sometimes expressed as the assumption that we are 'not special', or that we should expect to be 'average'[iv]. In my terminology, we can say that the epistemic process is random. This means

P(I'm person 1 | H1) = 1

P((I'm person 1 | H2) = P((I'm person 2 | H2) = 1/2

This puts us in the top row of the table.

The second question is: is the property you have found out about certain to be instantiated? Assuming you are person 1, the property is: the property of being person 1. This is certain to be instantiated given either H1 or H2. This puts us in the right column of the table. We can conclude that the Few Outcomes hypothesis, H1, is confirmed.

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No confirmation	(2) Few Outcomes confirmed **Doomsday Argument**
Persistent process	(2) Many Outcomes confirmed	(4) No confirmation

Table 3

An application of Bayes theorem confirms this:

E = I am person 1

$$P(H1 \mid E) = P(H1) * P(E \mid H1) / P(E)$$

$$= 1/2 * (1 / 3/4)$$

$$= 2/3$$

The evidence therefore confirms H1, which had a prior probability of only 1/2.. Intuitively, you are certain to be person 1 given H1, but there is only a 50% chance that you'll be person 1 given H2. So being person 1 confirms H1. The hypothesis that there are fewer people in the universe has been confirmed. This is the Doomsday argument.

There are many objections to the Doomsday Argument of course. It is not my purpose to discuss any of them here (see Richmond 2006 for a comprehensive review). My aim is to fit the Doomsday Argument into a more general taxonomy and hopefully to show by assimilation to other cases that no fallacies are made in the reasoning. If you reject one of the premises of the Doomsday Argument, you must be happy to reject the analogous premise of all the other arguments. The next case we examine also involves discovering a property that is certain to be instantiated, but with a different epistemic process.

Sleeping Beauty

If we change the process from random to persistent, and the self-locating variable from agent to time, we get the Sleeping Beauty problem (Elga 2000):

It is Sunday night. Sleeping Beauty is about to be drugged and put to sleep. She will be woken briefly on Monday. Then she will be put back to sleep and her memory of being awoken will be erased. She might be awoken on Tuesday. Whether or not she is depends on the result of the toss of a fair coin. If it lands Heads, she will not be woken. She will sleep straight through to Wednesday, and the experiment will be over. If it lands Tails, she will be awoken on Tuesday. The Monday and Tuesday awakenings will be indistinguishable. Sleeping Beauty knows the setup of the experiment and is a paragon of probabilistic rationality.

	Monday	**Tuesday**
Heads	Awake	(Asleep)
Tails	Awake	Awake

Figure 5: Sleeping Beauty

Does being awoken confirm Tails?

Some say no; her credence in Heads should stay at 1/2. Call these *Halfers.*

Some say yes; her credence in Heads should fall to 1/3. Call these *Thirders.*

There are two types of thirder. Some think she does not learn any new evidence on being woken (Elga *ibid.*). Others think that Beauty does learn new evidence on being woken, and updating on this new evidence confirms Tails (Weintraub 2004, Horgan 2007). I think this latter position, that Beauty learns new evidence, is based on a failure to correctly take the selection effects into account.

Recall our two selection process:

> Random: An outcome is picked at random from the actual outcomes. There is no bias in the selection process.
>
> Persistence: If an outcome with a particular property exists, the observer is told about that outcome. (If it doesn't exist, an outcome is picked at random.)

The outcomes are the days. The relevant property is the property of being a day on which Beauty is awake. Which process best models Beauty's learning of 'I'm awake'? Persistence. Beauty can only learn about days when she is awake. She does not randomly find herself on a day when she may or may not be awake. Rather, she is biased towards observing days when she is awake; she cannot observe days when she is asleep. Let's go through both processes.

Random Process

Assume you are Beauty. Suppose your being awake is discovered with a random process. This could happen if your observation of a day was independent of whether you were awake on that day. We could picture this as your time-traveling sub-conscious picking a day at random and peeking in to see if you're awake. If this is the process, and you turn out to be awake, Tails is confirmed and Heads disconfirmed, just as thirders say.

$$\text{P(Heads | I'm awake)} = \frac{\text{P(I'm awake \& Heads)}}{\text{P(I'm awake)}}$$

$$= \frac{1/4}{3/4} = \frac{1}{3}$$

Heads		**Tails**	
Monday	**Tuesday**	**Monday**	**Tuesday**
I'm awake	I'm not awake	I'm awake	I'm awake

Figure 3: Learning I'm Awake with a Random Process

But clearly this is not the process by which Beauty finds herself to be awake. Due to the nature of being awake, Beauty can only observe days when she is awake.

Persistent Process

As only waking days can be observed, there is a bias towards observing days on which there is an awakening.

$$\text{P(Heads | I'm awake)} = \frac{\text{P(I'm awake \& Heads)}}{\text{P(I'm awake)}}$$

$$= \frac{1/2}{1} = \frac{1}{2}$$

Heads	**Tails**
‘I am awake’	‘I am awake’

Figure 3: Learning ‘I’m Awake’ with a Persistent Process

It is certain that at least one day will have the property of being a day on which Beauty is awake, and there is a bias towards observing such days. So the result is no confirmation (box 4).

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No shift	(3) Few Outcomes confirmed Doomsday Argument
Persistent process	(2) Many Outcomes confirmed	(4) No shift **Sleeping Beauty**

Table 4

This answer coheres with the halfer position. There may be other arguments for being a thirder. But updating on 'I'm awake' does not give support to thirders. However, there is a similar argument in the vicinity where the Many Outcomes hypothesis *is* confirmed.

Generalized Sleeping Beauty

Let's modify the Sleeping Beauty story a little.

> A random waking device has an adjustable chance $c \in (0, 1]$ of waking Sleeping Beauty when activated on an occasion. In those circumstances in the original story where Beauty was awakened, we now suppose only that this waking device is activated. When $c = 1$, we have the original Sleeping Beauty problem. But if $c < 1$, the case is significantly different. (White 2006)

Whereas Sleeping Beauty has the structure of *Certain Aces,* Generalized Sleeping Beauty has a structure more similar to *Uncertain Aces*[v]. It is no longer certain that Beauty will be woken at all.

	Monday	Tuesday
Heads	Awake or asleep	(Asleep)
Tails	Awake or asleep	Awake or asleep

Figure 6: Sleeping Beauty

Beauty is now more likely to be woken given Tails than Heads. The reason is that there are two chances for Beauty to be woken given Heads, but only one given Tails.

Suppose Beauty finds herself awake. Does this confirm Tails? Yes. The epistemic process is still persistent. So we have a case where the process is persistent and the property (being a waking day) is not certain to be instantiated.

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No shift	(3) Few Outcomes confirmed Doomsday Argument
Persistent process	(2) Many Outcomes confirmed **Generalized Sleeping Beauty**	(4) No shift Sleeping Beauty

Table 5

Suppose the probability of Beauty being woken on each day is c. Call the proposition that Beauty is woken at least once during the experiment W. Beauty's new credence in H should be

$$P(H \mid W) = P(H)P(W \mid H)/[P(H)P(W \mid H) + P(-H)(P(W \mid -H)$$
$$= (1/2c)/[(1/2)c + (1/2)(1-(1-c)^2)]$$
$$= 1/(3-c)$$

As c → 1, P(H|W) → 1/2 (the halfer's position)

As c → 0, P(H|W) → 1/3 (the thirder's position)

White effectively uses Generalized Sleeping Beauty as an argument against thirders. I want to use it as a stepping stone to the third problem of self-locating evidence: the fine-tuning argument.

Fine-Tuning

If the fundamental constants of science had been much different from their actual values, life could not have existed in the universe. For example, if gravity had been a bit stronger, the universe would have collapsed in on itself moments after the big bang. If it had been a bit weaker, the universe would have flown apart so fast that molecules could never have been formed. The same holds for nearly all the other fundamental constants (see McMullin (1993)). (The initial conditions are also vital. For ease of exposition, I will take 'right constants' to include right initial conditions.) The existence of every living thing in the universe is balanced on a knife-edge[vi]. Nevertheless, life exists.

Proponents of the fine-tuning argument claim that the existence of life requires an explanation[vii]. One explanation is that there are many universes, and these universes have fundamental constants with different values. This is not David Lewis's theory that there are many possible worlds. By 'many universes' I mean many *actual* regions that are causally separated. This is a scientific hypothesis that has been independently suggested many times for various reasons, the latest being that the equations of string theory have approximately 100^{500} solutions, just one of which is instantiated by our universe. Perhaps all the other solutions are instantiated too.

We can model the simplest version of the fine-tuning argument as having two hypotheses, with either one or two universes.

	Self-Locating Possibility 1	Self-Locating Possibility 2
Universe	Right constants or wrong constants	-
Multiverse	Right constants or wrong constants	Right constants or wrong constants

Figure 7: Fine-tuning

Assume each universe has a certain chance, b, of having the right constants for life. Suppose we discover that some universe has the right constants for life. Does this confirm the Multiverse hypothesis? Yes. First of all, what is the epistemic process by which a universe with the right constants for life has been discovered? That is, if there is more than one universe, are we more likely to discover one with the right constants for life than the right constants for life? Well the process by which we observe a universe is by existing in one. It is impossible for us to observe a universe with the wrong constants for life, and there is a greater than zero probability that we will observe a universe with the right constants for life. So this gives us a bias towards discovering a universe with the right constants for life, which puts us in the bottom row.

Secondly, the property – being a universe with the right constants for life – was not certain to be instantiated. It was possible that no universe had the right constants for life. This puts us in the left column.

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No shift	(2) Few Outcomes confirmed Doomsday Argument
Persistent process	(2) Many Outcomes confirmed Generalized Sleeping Beauty **Fine-Tuning Argument**	(4) No shift Sleeping Beauty

Table 6

We can conclude that the discovery of life does confirm the Multiverse hypothesis. Let's go through the calculations. Let the probability of life in any given universe be b. Let L be the proposition that at least one universe has the right constants for life, and UV be the proposition that there is just one universe. Assume P(UV) = 1/2.

$$
\begin{aligned}
&P(UV \mid L) = P(UV)P(L \mid UV)/[P(UV)P(L \mid H) + P(-UV)(P(L \mid -UV) \\
&= (1/2b)/[(1/2)b + (1/2)(1-(1-b)^2)] \\
&= 1/(3-b)
\end{aligned}
$$

As b → 1, P(UV|L) → 1/2

As b → 0, P(UV|L) → 1/3

We can now dispense with an influential objection to this argument. Gould (1990) eloquently expresses the objection:

> Any complex historical outcome – intelligent life on earth for example – represents a summation of improbabilities and becomes therefore absurdly unlikely. But something has to happen, even if any particular "something" must stun us by its improbability. We could look at any particular outcome and say "Ain't it amazing. If the laws of nature had been set up a tad differently, we couldn't have this kind of universe at all". (p.183)

The same point can be found in Hacking (1987), Carlson and Olsson (1998), White (2000) and Juhl (2005). This objection assumes that we are observing some randomly selected universe. If we were, we would be in box 1, and the objection would be correct. But we are not observing a randomly selected universe that just happens to have the right constants for life. Instead, we could only have observed a universe with the right constants for life. The evidence is persistent, and so it is more likely to be found given Many Universes. The fourth and final problem case is also one in which the property we discover is not certain to be instantiated. But in this case there is no bias in the process.

Many Worlds

Stochastic quantum mechanics (QMS) is an indeterminstic theory. It says that when a measurement is made, there is a certain chance of each possible outcome occurring. Everettian quantum mechanics (QME) is a deterministic theory. It says that when a measurement is made, the universe divides, and there is a branch in which each outcome occurs. If we don't know which of these is true, and we are about to measure the spin of a particle, we get the following structure:

	Branch 1	**Branch 2**
QMS	Up or Down	-
QME	Up	Down

Figure 8: Quantum mechanics

According QMS, there is only one branch. One outcome will occur, which in this case is either Up or Down. The result of an observation will tell you which. According to QME, there are two branches. Making an observation tells you which branch you are on – the Up branch or the Down branch.

Suppose you make an observation and observe 'Up'. E = 'Up' is observed. Does this evidence confirm QME or QMS? First we must ask about the epistemic process. By what process was 'Up' discovered? Was there a bias towards observing 'Up'? That is, if QME is true, are you more likely to observe 'Up' than 'Down'? No. There is no bias towards one rather than the other. (This not to say that you are equally likely to see 'Up' as 'Down'. The branches may be of unequal weights. The point is that you are not more likely to be on the Up branch in virtue of it being the Up branch. Compare: you are more likely to observe a universe in virtue of it having right constants for life – true; you are more likely to observe an ‘Up’ branch in virtue of it being an ‘Up’ branch – false.) The process is random, so this puts as in the top row.

Secondly, is it certain that the ‘Up’ result occurs at all? Is the property of being an Up branch certain to be instantiated? No. If QMS is true, and if Down occurs, there is no Up branch. This puts us in the left column. So we get the result that there is no confirmation of either QME or QMS when ‘Up’ is observed.

	p is not certain to be instantiated	p is certain to be instantiated
Random process	(1) No shift **Quantum Mechanics**	(2) Few Outcomes confirmed Doomsday Argument
Persistent process	(2) Many Outcomes confirmed Generalized Sleeping Beauty Fine-Tuning Argument	(4) No shift Sleeping Beauty

Table 7

This result is what we wanted. It would have been embarrassing if an observation confirmed one of the interpretations over another. But there is a tempting line of thought that takes us down this path. That is, it is tempting to think that QME is confirmed by any observation. After all, 'Up' has a 100% chance of being observed if QME is true, but a lower chance if QMS is true. The same is true of 'Down'. It is a theorem that if H entails E, and E has a non-extreme prior probability, then E confirms H. So it appears that the observation of 'Up' (or 'Down') must confirm QME. This kind of reasoning is called 'Naïve Conditionalization' by Greaves (2006), who wants nothing to do with it. But Price (2006) is tempted, as are Titelbaum (ms), Dorr (ms) in the context of Sleeping Beauty[viii]. I think they make the mistake of not paying attention to the process. It is true that 'Up' has been observed, and that 'Up' is more likely to be observed given QME. But Price at al. implicitly assume that there is a bias towards discovering 'Up' (or 'Down', or whatever the result happens to be). If there were such a bias, then discovering 'Up' would indeed confirm QME. But there is no such bias. The process is random, so the observation does not confirm QME. The total evidence is E* = 'I observed 'Up' due to a random process'[ix]. This is not entailed by QME.

Conclusion

The relevance of the selection process to confirmation is easily over-looked. Far from being an unusual feature of certain cases, selection effects are ubiquitous. Even if they don't cause any confusion, they are still there. Just as Frege claimed we can't have a reference without a sense, so we can't have evidence without a process. Problems arise when self-locating evidence is involved because condition (U), that there is only one actual outcome, tends to fail. Much of the disagreement concerning self-locating evidence has nothing to do with self-locating evidence *per se*. The disagreement comes from failing to take the selection effects into account. Once we do this, the landscape looks much simpler. Four problems with very different subject matter can be seen to have the same structure. The differences come from the selection process and whether the property discovered was certain to be instantiated. I have defended four arguments – the Doomsday Argument, the halfer position in Sleeping Beauty, the Fine-tuning argument, and the applicability of confirmation theory to Everett interpretations of quantum mechanics. Anyone who wanted to endorse one of these arguments and not another would have to demonstrate the disanalogy between the cases.

References

Bradley, D. and Fitelson. B (2003) 'Monty Hall, Doomsday and Confirmation' *Analysis* 63: 23-31

Bostrom, N. (2002) *Anthropic Bias: Observation Selection Effects in Science and Philosophy.* New York : Routledge**.**

Carlson, E. and Olsson, E. J. (1998) 'Is Our Existence in Need of Further Explanation?' *Inquiry* 41: 283-309.

Carter, B. (1974) 'Large Number Coincidences and the Anthropic Principle in cosmology' *Confrontation of Cosmological Theories with Data* M. S. Longair. Dordrecht, Reidel: 291-8.

Craig, W. L. (1988) *'Barrow and Tipler on the Anthropic Principle vs. Design' British Journal for the Philosophy of Science 38: 389-395.*

Dorr, C. (ms) 'A Challenge for Halfers'

Elga, Adam (2000), ‘Self-locating belief and the Sleeping Beauty problem’, *Analysis* 60: 143–147.

– (2004) ‘Defeating Dr. Evil with Self-Locating Belief’ *Philosophy and Phenomenological Research* 69(2).

Everett, H. III (1957) “Relative State’ formulation of quantum mechanics’ *Reviews of Modern Physics* 35: 423-456

Greaves, H. (2004) ‘Understanding Deutsch’s Probability in a Deterministic Universe’. *Studies in History and Philosophy of Modern Physics.*

- (2006) ‘Probability in the Everett Interpretation: A Solution to the Epistemic Problem’ (ms)

Hacking, I. (1987) ‘The Inverse Gambler’s Fallacy: The Argument from Design. The Anthropic Principle Applied to Wheeler Universes’ *Mind* 76: 331-340.

Horgan, T. (2004) ‘Sleeping Beauty Awakened: New Odds at the Dawn of the New Day’ *Analysis* 64: 10-21

– (2007) ‘Synchronic Bayesian updating and the generalized Sleeping Beauty problem’ *Analysis* 67:293, 50–59

Hutchison, K. (1999) ‘What Are Conditional Probabilities Conditional On?’ *British Journal for the Philosophy of Science* 50, 665-95.

Juhl, C. (2005) ‘Fine‐tuning, Many Worlds, and the 'Inverse Gambler's Fallacy'’ *Nous* 39 (2): 337–347.

McMullin, E. (1993) ‘Indifference Principle and Anthropic Principle in Cosmology’, *Studies in the History and Philosophy of Science*, 24: 359–89.

Meacham, C. (forthcoming) 'Sleeping Beauty and the Dynamics of *De Se* Belief' *Philosophical Studies*

Lewis, D. (2001) "Sleeping Beauty: Reply to Elga", *Analysis* 61: 171–176.

Lewis, P. (forthcoming) 'Quantum Sleeping Beauty' *Analysis.*

Leslie, J. (1986) *Universes.* London: Routledge

– (1996) *The End of the World: The Science and Ethics of Human Extinction.* London: Routledge

Parfit, D. (1998) *Why anything? Why this?* London Review of Books, 1998. Jan 22, 24-27.

Price, Huw (2006) 'Probability in the Everett World: Comments on Wallace and Greaves'. Available online from http://philsci-archive.pitt.edu/archive/00002719/

Swinburne, R. (1990) 'Arguments from the Fine-Tuning of the Universe', in J. Leslie (ed.), *Physical Cosmology and Philosophy*, New York: MacMillan, 160–79.

Titelbaum, M. (ms) 'When are Context-Sensitive Beliefs Relevant?'

van Inwagen, P. (1993) *Metaphysics*, Boulder, CO: Westview Press.

Weintraub, R. (2004) 'Sleeping Beauty: A Simple Solution' *Analysis* 64: 8-10

White, R. (2000) 'Fine-tuning and Multiple Universes' *Nous* 34 (2): 260-276

-- (forthcoming) 'The generalized Sleeping Beauty problem: A challenge for thirders', *Analysis*.

---

[i] Randomness is assumed by statisticians and persistence is assumed by Bayesians. Both are very natural assumptions, and I think this goes some way to explaining how easy it is to over-look that a substantive assumption has been made in both cases.

[ii] I would like to say that a piece of evidence is a proposition, but ‘proposition’ has the connotation of ‘uncentred proposition’, so I cannot.

[iii] See Carter 1974, Leslie 1989, 1996

[iv] This assumption is given a rigorous defence by Elga 2004. Though see Bradley and Fitelson 2001 for a relaxation of this assumption in the context of the Doomsday Argument.

[v] There is a slight difference between Generalized Sleeping Beauty and Uncertain Aces, in which both an Ace and a King are dealt if Tails. The difference is that even if Tails lands, Beauty might not be woken. If Generalized Sleeping Beauty had the same structure as *Uncertain Aces*, then Beauty would be woken on Monday and sleep on Tuesday (or the other way round). But this difference only affects the degree of confirmation, not the direction.

[vi] I have found that some philosophers are hostile to this claim for reasons I find puzzling; physicists seem to take the claim as data. I am not especially concerned with the truth of the claim however, but with the inferences that would be valid if it were true.

[vii] These proponents include Parfit 1998, Leslie 1989, Swinburne 1991, Craig 1988 and van Inwagen 1993.

[viii] The close connection between Sleeping Beauty and the Everett interpretation is pointed out by Peter Lewis (forthcoming).

[ix] It is vital to notice that the ‘I’ refers to the observer-part on just one of the branches after division. QME does entail that I – an entity that exists in *both* branches – observe Up. But it does not entail that I – the observer-part on one of the later branches – observe Up. It is an the observer-part on a particular branch that makes an observation. Epistemically speaking, the observer-parts are on their own.