In Defense of Statistical Evidence: From Epistemology to Courtrooms

By Kamyar Asasi

Submitted to

Central European University

Department of Philosophy

In partial fulfilment of the requirements for the degree of Master of Arts

Supervisor: Professor Timothy Martin Crane

BUDAPEST, HUNGARY
JUNE 2019

TABLE OF CONTENTS

INTRODUCTION

Imagine you buy a ticket for a lottery with 10 million tickets and only one winner. Now, suppose you buy a ticket in a similar lottery every single day. What is your first impression of the amount of time that it takes until you win once? A couple of years? In fact, the most probable outcome would be winning the lottery after 27,000 years! It means if you wanted to have a good chance of winning at least once until now, you should have started buying your daily tickets from prehistory. Even in that case, assuming the price of each ticket is $1 and the prize at best $10 million, it is most probable that you (in fact, your next generations) have lost interest of more than $1 billion given a 1% interest rate (the more relevant compound interest exceeds a line of digits).

Accordingly, forming a belief like "my ticket is a loser" based on such statistical evidence seems very rational. After all, if I and all my friends (and even our several next generations) constantly form such beliefs in our lifetime, it is most likely that none of them would be false. Moreover, statistical evidence does not seem to be inferior to non-statistical evidence in terms of providing justification; this is because as Smith (2018, p. 1198, emphasis in original) puts it: "Though they differ over the details, many epistemologists agree that securing justification for believing a proposition is a matter of ensuring that it is sufficiently *probable*, given one's evidence. Such a view offers no explanation as to why we would privilege testimonial over statistical evidence. Both kinds of evidence are equally capable of making propositions probable[1]".

[1] Later in that paper, Smith argues that the problem with statistical evidence is that it, unlike individualized evidence, does not provide normic support. I will get to this in chapter 3.

However, on the one hand. many epistemologists assume that statistical evidence cannot be an acceptable basis for knowledge like in the aforementioned lottery case or the following famous version of it:

> *Lottie*
>
> Lottie has a ticket for a fair lottery with very long odds. The lottery has been drawn, although Lottie has not heard the result yet. Reflecting on the odds involved she concludes that her ticket is a loser. Lottie's belief that she owns a losing ticket is true (Pritchard, 2012, p. 252)[2].

On the other hand, people doubt that courts can base convictions on statistical evidence such as the following:

> *Joe and television*
>
> An electronics store is struck by looters during a riot – 100 people walk out of the store carrying televisions, while the transaction record at the cash register indicates that only one television was paid for, though no receipt was issued. Suppose Joe is apprehended carrying a television from the store, but we have no other information about him (Smith, 2018, p.1196).

Despite the extreme amount of reliability that statistical evidence can provide, many of us find the idea of basing knowledge or conviction (in a court of law) on statistical evidence counter-intuitive. These intuitions hold even when we are aware of the fact that in some cases with non-statistical evidence (like eyewitness testimony) where we are happy to attribute knowledge, the chances of forming a false belief (or the probabilistic likelihood of error) are much higher than the chances of forming a false belief in many cases with statistical evidence.

In my thesis, I will argue that, in principle, statistical evidence can be a basis for knowledge and conviction in a court of law, and there is no rationale for discriminating against statistical evidence; the intuitive difference between non-statistical and statistical

[2] In that paper, Pritchard does not directly discuss the problem with statistical evidence, but such problem with respect to similar lottery cases has been discussed by many like Thomson, 1986; Smith, 2018; Enoch et al., 2012; etc.

evidence is merely psychological and can be understood in the light of consideration like the historical development of our intuitions.

I start with epistemology. In chapter 1, I suggest some scenarios in which the belief is based on statistical evidence yet, according to the criteria from accounts of knowledge like causal theory (Goldman, 1967), Sensitivity (Nozick, 1981), and Safety (Pritchard, 2012), in the market for knowledge. In chapter 2, I apply a famous distinction from Wilfrid Sellars (1963) between 'manifest image' and 'scientific image' of the world and ourselves to the case of knowledge; I enumerate some of the problems with the intuitive notion of knowledge, discuss some historical considerations to better understand such intuitive reactions and, briefly, explore the possibility of a move towards a more scientific notion of knowledge which can help with addressing some of those issues.

In chapter 3, I turn to a parallel issue with statistical evidence in the context of legal cases. Discussing two proposals in the literature from Judith Thomson (1986) and Martin Smith (2018) and a third possible strategy to distinguish the non-statistical and statistical evidence, I argue that none of them manages to provide an independent rationale for such discrimination.

Finally, in chapter 4, I address some of the worries about the reliability of the current statistical methods that are used to produce statistical evidence. Then, I discuss the applicability of statistical methods to particular cases and propose some criteria to decide whether the statistical evidence is relevant and can be used as a basis for judgment in each scenario.

I. CHAPTER 1: STATISTICAL EVIDENCE IN EPISTEMOLOGY

I.1 INTRODUCTION

Many epistemologists hold that Lottie does not have knowledge. However, the belief in cases like Lottie is not only based on statistical evidence but also insensitive, unsafe, and not causally connected with the fact which I suspect boosts the counter-intuitiveness of knowledge assignment to such cases. Note that, the fact that the belief is based on statistical evidence does not always go hand in hand with it being unsafe, insensitive or not causally connected (i.e. not being a candidate for knowledge), and these two aspects can vary quite independently. In what follows, I suggest some cases in which the belief is based on statistical evidence, yet they satisfy the criteria of such accounts of knowledge and, thus, can be considered as counter-examples for the claim that knowledge cannot be based on statistical evidence.

I.2 A CASE FOR KNOWLEDGE BASED ON STATISTICAL EVIDENCE

Consider the following scenarios:

> **1st Scenario:** Arash is in a park and knows that there are exactly 1 million animals, all of which are real sheep. One of the animals enters his visual field and he forms the belief that there is a sheep in front of him. It is indeed a sheep.

> **2nd Scenario:** Arash goes to another park and knows that everything is similar to the 1st scenario with the exception that one of these 1 million animals is a dog carefully disguised as a sheep. Again, one random animal enters his visual field, and he forms the belief that there is a sheep in front of him. It is indeed a sheep (suppose replacing this animal with the disguised dog is not easy: The park is arbitrarily large, say the size of a continent, and, as Arash can see, the animal in front of him is the only animal around, say, the next animal is thousands of kilometers away; this animal and the disguised dog have never been in close proximity in the past; sheep-allocation mechanism does not allow for easy counterfactual replacement; etc.).

The 1st scenario is an ordinary case of knowledge based on perception while Arash's belief in the 2nd scenario involves statistical evidence. We have a clear intuition about the knowledge status of his belief in the 1st scenario, yet people's intuitions conflict when it comes to knowledge assignment in the 2nd scenario. I will argue that, despite it being counter-intuitive for some people, according to many theories of knowledge the belief in the 2nd scenario can be considered a case of knowledge. I will start by Goldman's famous causal theory as an anti-luck predecessor of reliabilism and continue to more recent reliabilist theories: sensitivity and safety accounts.

I.2.1 A Causal Theory of Knowledge

The causal theory of knowledge was proposed by Alvin Goldman in 1967. He suggested replacing the justification clause in the classical analysis of knowledge as justified true belief (JTB) with a condition that requires a causal connection between the belief and the fact believed. According to him:

> *S knows that p* if and only if
> *the fact p is causally connected in an "appropriate" way with S's believing p.*
> (Goldman, 1967, p. 369, emphasis in original)

He explains that "appropriate" causal processes include: perception; memory; a causal chain that exemplifies memory or perception and is correctly reconstructed by inferences; and combinations of these (Goldman, 1967).

Many think that this account is too demanding and excludes many cases that we naturally take as cases of knowledge; moreover, although this analysis was originally motivated by Gettier cases, it fails to respond to many standard counterexamples to JTB theory. So, despite being intuitive, this theory is not very popular (Goldman himself dropped this theory later). In any case, it is obvious that the case described in the 2nd scenario fulfills the requirements for a causal theory of knowledge: The fact that there is a sheep there causes Arash, through perception, to believe that there is a sheep in front of him. Thus, as far as the causal theory is concerned, Arash's belief in the 2nd scenario, despite being based on statistical evidence, is a candidate for knowledge.

I.2.2 Sensitivity

According to this account adding the *sensitivity* condition on knowledge does the trick:

> **Sensitivity:** S's belief that p is sensitive if and only if, if p were false, S would not believe that p.

Given a Lewisian semantics for counterfactual conditionals (Lewis, 1973), this condition is equivalent to saying that in the nearest possible world in which p is false, the subject does not believe that p (Nozick, 1981).

Applying it to the 2nd scenario, Arash's belief is sensitive since in the nearest possible world in which his belief is false, namely, there is no sheep in front of him, he wouldn't believe so. This conclusion is not a surprise given, as the previous discussion of the causal theory showed, his belief has been caused by the fact that there is a sheep in front of him; so, had there wasn't a sheep, he wouldn't believe so.

One may object that in a possible world in which the disguised dog is in the place of the sheep Arash would still form the belief that there is a sheep in front of him and thus his belief is not sensitive. In response, note that for many cases that everyone thinks are sensitive one may still find a possible world in which the belief is false but the subject still believes it; so, crucially in order to test the sensitivity condition, we need to consider the *closest possible worlds* in which subject's belief is false. Regarding the 2nd scenario, the relevant *closest* possible worlds seem to be worlds in which there is no sheep in front of him (say a small change in the direction of the sheep's walk, prevents it from entering Arash's visual field) and therefore he does not believe so, not a possible world in which the disguised dog has been brought to replace the sheep in Arash's sight (say the direction and speed of both the sheep and the disguised dog have been changed in a way to replace each other).

I.2.3 Safety

The previous sensitivity account is not very popular. One serious problem with the sensitivity condition is that it is too demanding and cannot accommodate many ordinary cases that we tend to take as knowledge like the following case of inductive knowledge:

> **Ernie:** Ernie deposits a rubbish bag into the rubbish chute in his high-rise flat. He has every reason to think that the chute is working correctly and so believes, a few minutes later, that the bag is in the basement. His belief is true. (Pritchard repeats this example from Sosa 1999; in Pritchard 2012, p. 253).

Ernie's belief seems an ordinary case of knowledge, however, as Pritchard (2012) correctly argues, it is not sensitive: Even if there had been some problems with the chute that had prevented the bag from reaching the basement, he would still have believed the bag was there. Such problems have made accounts like safety much more popular than sensitivity.

Pritchard (2012) characterizes *safety* as follows:

> **Safety:** If S knows that p then S's true belief that p could not easily have been false (p. 253).

He argues that the analysis of knowledge in this way can account for cases like Ernie: Ernie knows that the bag is in the basement because, considering the conscientious way that he uses to form his belief, his belief could not easily have been false and therefore is safe (Pritchard, 2012).

Pritchard (2012) explains that "could not easily" clause in the criteria for safety has to do with modal closeness, namely the amount of change required to the actual circumstances to make the belief false. He explains that safety does not require immunity from any possibility of error no matter how remote, rather demands immunity from the belief becoming false with very small changes in the actual world (namely, errors that can happen with very small changes in actual circumstances).

On reflection, in the 2nd scenario, Arash's belief is safe. This is because in near-by possible worlds when he forms the belief, the belief is true. Obviously, there are also possible worlds in which there is the disguised dog in place of the sheep and he forms a false belief, but such possible worlds are not very close: Such possible worlds not only demand a change in the direction of the sheep which came to Arash's sight but also—in order to replace it with the disguised dog—require changes in the moving direction, the speed, etc., of the dog which can be anywhere in the park (since there are 1 million animals spread in the park, it is very unlikely that the dog is one of the sheep's neighboring animals, but even if it is the closest animal to the sheep, it would still be thousands of kilometers away).

If one rejects the safety of Arash's belief in the 2nd scenario, she will be pushed to reject the safety of Arash's belief in the following scenario as well:

3rd scenario: Suppose Arash knows the following:

- A while ago, the country imported 2 million animals: 1,999,999 sheep and 1 dog which is carefully disguised as a sheep;
- These animals have been randomly distributed between two main national parks (1 million animals in each national park).

Arash goes to one of these parks (in which, unbeknown to him, all the animals are in fact real sheep, namely the disguised dog is in the other park) and faces a sheep-like animal (which is indeed a sheep) and forms the belief that it is a sheep.

Many—including many of those who don't assign knowledge in the 2nd scenario—tend to think that Arash has knowledge in the 3rd scenario; all of the sheep in the national park are real after all. Nevertheless, one may argue that, in the 3rd scenario, with an amount of change which is not bigger than the required change to make Arash's belief false in the 2nd scenario back in the time of distributing animals between the national parks, a disguised dog could have replaced the animal that Arash is looking at, and therefore his belief is not safe.

More importantly, there is a worry that denial of knowledge assignment in the 2nd scenario may lead to skepticism about many ordinary cases of knowledge in which the belief could undetectably become false by making the same (or even smaller) amount of change in the actual world compared with the 2nd scenario. For instance, consider someone sees a real sheep in a meadow and forms the belief that there is a sheep in front of him. Now it seems very odd to say that just by adding to the story that there is a dog disguised as a sheep in that country, say thousands of kilometers away in the other end of the meadow which is in another province, he doesn't know that there is a sheep in front of him anymore. Note that, in this scenario, the required amount of change to make the belief false is not less than that of the 2nd scenario; in both cases, we need to replace the real sheep with a faraway dog.

I.3 A Case for Knowledge Based on "Merely Statistical" Evidence

The previous 2nd scenario can be seen as analogous to a lottery case. One million animals are like one million tickets; each of 999,999 sheep is like a losing ticket and the disguised dog is the winning one. Accordingly, if one concedes that in the 2nd scenario Arash's belief is safe, and he has knowledge (despite its counter-intuitiveness for some people), she should admit that the same is true of such lottery case.

However, one may object that in ordinary lotteries, unlike the 2nd scenario, there is no causal connection between the fact that the ticket is a loser and the subject's belief. Therefore, she may argue that the presence of a causal connection between the fact that there is a sheep there and the corresponding belief in the 2nd scenario is what allows knowledge assignment. Consequently, she may accept that Arash has knowledge in the 2nd scenario but still refrain from attributing knowledge to cases in which there is no causal connection between the belief and the fact and the subject forms the belief by solely reflecting on the odds, claiming that such cases with "merely (purely/solely) statistical" evidence (Thomson, 1986; Smith, 2018; Enoch et al., 2012) are, in fact, the ones that epistemologists are interested in when they discuss statistical evidence.

In reply, considering that the dog is disguised as a sheep and thus the mere fact that the animal looks like a sheep to Arash does not provide any evidence for it being a real sheep, I think even in the 2nd scenario his belief that the animal is a sheep is merely supported by the statistical data. However, one may object that the real story is that Arash first forms an ordinary perceptual belief that there is a sheep in front of him after seeing the animal, then the statistical data comes in and may or may not defeat this belief.

In order to avoid such complications and objections, I've devised the following scenario in a way that blocks the direct causal connection between the fact and the belief (hence avoids the worries about the subject's evidence not being merely statistical):

> **4th scenario:** Suppose Arash knows that there are 1 million animals in the park: 999,999 sheep and 1 dog carefully disguised as a sheep in a random place among them (just like the 2nd scenario). Suppose one of these 1 million animals is behind a wall in front of him and he knows it (say, someone tells him so, or he hears an animal-like sound, etc.). Reflecting on the odds involved, he concludes that the animal behind the wall is a sheep. His belief is true.

The 4th scenario is very similar to an ordinary lottery: Imagine someone finds a lottery ticket in the street after the lottery has been drawn (she knows that one of 1 million animals in the park is behind the wall), reflecting on the odds involved, she concludes that the ticket is a loser (the animal behind the wall is a sheep). The ticket is, in fact, a losing one (the animal is, in fact, a sheep).

Needless to say, as noted earlier, an analysis of knowledge in terms of causation is too demanding, and therefore the mere fact that a belief is not causally connected with the fact does not imply that it cannot be a case of knowledge. That said, as far as safety (and arguably sensitivity) is concerned, the 4th scenario is very similar to the 2nd scenario (i.e. a possible world in which the disguised dog has replaced the sheep is not close, etc.) and, thus, Arash's belief in the 4th scenario can be a candidate for knowledge.

One may object that, even in this case, Arash's belief is not based on merely statistical evidence since it still includes some causal connections; for example, someone may have told

Arash that one of these animals is behind the wall or that there are 1 million animals in the park, etc. In reply, note that when people talk about merely statistical evidence, they don't mean that all the pieces in the body of evidence behind a particular conclusion are merely statistical, rather a specific piece of evidence which is crucial for drawing the conclusion is merely statistical (that said, I think the literature will be benefited from more discussions on this point). For instance, even in Lottie case the body of evidence, in addition to her merely statistical evidence, includes many pieces of non-statistical evidence like (approximate) number of all tickets as well as number of winning tickets in the lottery, the evidence that she owns a ticket (which corresponds to the evidence in the 4th scenario that one of those 1 million animals is behind the wall), etc., without which she wouldn't be able to draw the conclusion that she has a losing ticket.

I.3.1 Responding to A Worry About Gettier-Style Cases

One worry that may arise from allowing knowledge assignment to statistical cases like the 4th scenario is that it may lead to allow knowledge assignment to Gettier-style cases as well. However, note that, although safety account is designed to respond to problems posed by Gettier-style cases as well as lottery cases, these problems are distinct and allowing knowledge assignment to some lottery-style cases does not necessarily commit us to allowing it for Gettier cases. To see how it is the case consider the following scenario which is a modified version of the 4th scenario that poses a lottery-style and a Gettier-style problem at the same time:

> **Gettiered 4th Scenario:** Suppose Arash knows that there are 1 million animals in the park: 999,999 sheep and 1 dog carefully disguised as a sheep in a random place among them. Suppose one of these 1 million animals is behind a wall in front of him and he knows it. Reflecting on the odds involved, he concludes that there is a sheep behind the wall. In fact, that animal is a disguised dog, but one meter away from it there is another animal behind the wall which Arash didn't know about. This animal is a real sheep and therefore Arash's belief that there is a sheep behind the wall is true.

Before getting to this case, recall that Arash's belief in the original 4th scenario is safe because a possible world in which Arash's belief undetectably is false is remote since it requires two changes: One is to remove the real sheep from behind the wall, and the other is to bring the disguised dog (which may be anywhere in the park) to replace it. Now, this, apparently, does not imply that Arash's belief in the Gettiered 4th scenario is safe too. Arash's belief in the Gettiered 4th scenario is, in fact, unsafe. This is because to make Arash's belief undetectably false in this scenario we don't need to make big changes from the actual world like bringing a faraway animal to replace another one. All is needed is a small change in the direction of the real sheep which places it somewhere other than behind the wall (even a relocation of a few meters to the front or beside the wall). Therefore, in the Gettiered 4th scenario, unlike the 4th scenario, the possible world in which Arash's belief is undetectably false is modally close to the actual world, and therefore his belief in this Gettier-style case is unsafe and not in the market for knowledge.

It is worth noting that even if someone comes up with a similar Gettier-style case in which the belief is safe, it would be a problem for the safety account in general and does not have much to do with the previous discussions to allow knowledge assignment in some statistical cases. In fact, in order to deal with a similar worry that safety condition alone cannot respond to many Gettier-style cases, Pritchard changed his view from an "Anti-luck Epistemology" (2007) which mainly relies on safety in responding to Gettier-style cases to an "Anti-luck Virtue Epistemology" (2012) which, in addition to safety, posits the extra requirement that one's cognitive success should be the product of one's relevant cognitive abilities.

Note that this new condition does not affect knowledge assignment in the previous scenarios. This is because the cognitive success in those scenarios is the product of Arash's cognitive ability in reflecting on the odds. The cognitive success would not be the result of

Arash's relevant cognitive abilities, say, in a reverse scenario in which there are 999,999 disguised dogs and only one sheep in the park, but Arash still forms the belief that the animal is a sheep. Pritchard himself argues that even Lottie's cognitive success is the product of exercising her relevant cognitive abilities, though, according to Pritchard, Lottie does not have knowledge for another reason namely since her belief is not safe (Pritchard, 2012). Accordingly, while this new ability condition leaves the knowledge assignment to the previous scenarios intact, it comes in handy in ruling out many Gettier-style cases. For example, one may argue that in the Gettiered 4th scenario, Arash's cognitive abilities is not responsible for the success, instead the success is due to the fact that (unbeknown to Arash) there is a second animal behind the wall which is a real sheep, without which Arash's belief would be undetectably false.

I.4 CONCLUSION

In this chapter, I have argued that a belief can be based on statistical evidence yet in cases like the 2nd scenario causally connected to the fact, sensitive, and safe, or in cases like the 4th scenario—which better resemble ordinary lotteries—safe and (arguably) sensitive and therefore in the market for knowledge.

This shows that the culpable behind the denial of safety and thus knowledge assignment in cases like Lottie is not their statistical aspect, rather the contingent facts about those specific situations that make the truth of subject's belief very sensitive to small counterfactual changes from the actual circumstance and, thus, unsafe. For instance, in Lottie case, what we might call the chaotic aspect of the process—which makes the result very sensitive to minor changes, say, in the movement of balls in the lottery machine—is responsible for making the belief unsafe and therefore not a candidate for knowledge. In contrast, the process in the 4th scenario is not similarly chaotic, and there are no small counterfactual changes that could have made Arash’s belief false which, in turn, makes the belief safe and, therefore, a case of knowledge despite the fact that it is as relied on statistical evidence as Lottie’s.

II. CHAPTER 2: MOVING TOWARDS A SCIENTIFIC IMAGE OF KNOWLEDGE

II.1 INTRODUCTION

I argued, in the previous chapter, that a belief can be based on statistical evidence, yet satisfy the criteria from many accounts of knowledge. However, a passionate proponent of the view that statistical evidence cannot be a basis for knowledge may react to previous cases differently by rejecting the plausibility of those accounts of knowledge like the safety that allow knowledge assignment in such scenarios. But where does this strong confidence in rejecting statistical evidence come from?

The main—and arguably the only—answer to this question is in reference to the intuitions. Some of us are intuitively uncomfortable with assigning knowledge in cases like the previous 4^{th} scenario, though we may not be able to provide any satisfactory rationale for that. Such intuitions are even stronger and are shared by more people when it comes to cases like Lottie, but even in those cases, we are short of adequate justifications—independent from the intuitions themselves—for our intuitive judgments. This apparent disparity between the intuitive picture, on the one hand, and a rigorous system of supporting arguments, on the other, may remind us of a famous distinction that Wilfrid Sellars (1963) makes between ‘manifest image’ and ‘scientific image’ we have of the world and ourselves.

Sellars characterizes the manifest image as the common sense and ordinary way we observe and explain the world, as the world appears to us and as we experience it. The scientific image, on the other hand, is often more rigorous and—as he stipulates—involves postulation of nonmanifest entities (he, in fact, suggests scientific image might better be called a 'postulational' or 'theoretical' image). He explains that the contrast between these images is not between an arbitrary, uncritical, and naïve conception, on the one hand, and a reflected, disciplined and 'scientific' on the other. He ultimately argues for a 'synoptic vision' and a 'stereoscopic view' that unites two pictures, though he seems to give the 'primacy', at the end of the day, to scientific image (Sellars, 1963).

In this chapter, I try to apply a similar distinction to the case of knowledge. Sellars distinction provides a worthwhile historical context and a good place to start, yet I won't completely commit myself to the way Sellars stipulates these images. Here, by 'manifest image' of knowledge I mean the common sense and intuitive notion of knowledge, while by 'scientific image' I refer to more systematic approaches which may not always fully conform to the intuitions and have become possible by relatively recent developments in science and mathematics.

In what follows, after discussing some of the problems for the manifest image, I suggest some historical considerations to make sense of it and, finally, briefly allude to some possibilities for a more scientific image of knowledge.

II.2 Some Problems with the Manifest Image of Knowledge

II.2.1 Historical Failure of Attempts to Capture the Manifest Image

For a long time, epistemologists have tried hard to provide a satisfactory analysis of knowledge, but these attempts have not been very successful. In response, some have rejected the coherence of the notion of knowledge and asked to abandon it altogether and replace it with a less problematic concept like true belief (see, for example, Papineau, forthcoming), and many others have decided to follow Williamson's famous knowledge-first account which rejects the project of analyzing knowledge (Williamson, 2000). In this context, one may argue that the strong commitments to fully capture the manifest image of knowledge in a theory of knowledge is responsible for the failure of attempts to provide a plausible analysis of it.

II.2.2 Problematic Emphasis on Modal Closeness

One main motivation for relying on modal closeness in accounts like sensitivity and safety can be traced back to our putative psychological tendency to care about the modal closeness of the error instead of the probability of it when it comes to knowledge ascription which seems to be in line with some empirical studies on people's judgments about risk and luck: For example, people seem to assign much more risk to a modally close event than to a modally remote one, even if they are fully aware that both events have the same probability of happening (Pritchard, 2012 mentions this from Pritchard and Smith, 2004).

Such risk assignment is understandable but—in the light of relatively recent tools like probability theory—not necessarily rational. The relation between modal closeness and probability of happening is not a simple one; in many cases they seem to go hand in hand, but in situations in which they diverge (or it seems to us that they diverge), it is expected from a rational agent (who cares about, say, avoiding some danger) to focus on the probability of happening and not how much it seems modally close. For instance, take the following case:

> *Gottie*
>
> Gottie has a ticket for a fair lottery with very long odds. The lottery machine, in this case, is similar to Lottie case in every aspect with the exception that it is much bigger. It is a gigantic machine with balls the size of a car (but the chance of winning/losing is the same). The lottery has been drawn, although Gottie has not heard the result yet. Reflecting on the odds involved he concludes that his ticket is a loser. Gottie's belief that he owns a losing ticket is true.

Although when it comes to probability of error Lottie and Gottie are in a similar situation, an account of knowledge like safety that concerns with modal closeness seems to treat them differently. This is because in Lottie case *very small* changes in the movement of the balls can make the belief false, whereas in Gottie case the required change to move those huge balls does not seem small. Therefore, Lottie's belief is unsafe, while Gottie's is safe (in fact, Gottie case can be seen as another example of knowledge based on statistical evidence).

Such different treatment of Lottie and Gottie cases by accounts like safety has some unfavorable implications: Firstly, it seems very odd to decide on whether one has knowledge in a lottery case only after inspecting the size of the lottery machine! After all, both lottery machines have the same functional role in producing random results, both lotteries are fair, the chance of winning/losing is the same, the evidence in both cases is similarly reliable, etc. So, why does the size of the machines should be relevant for knowledge assignment?

Secondly, claiming that Gottie has but Lottie does not have knowledge seems to be at odds with the fact that a rational agent acts similarly in both cases when it comes to producing decisions or actions (e.g. rejecting an offer to book an expensive trip that she can only afford if she wins the lottery).

To raise another problem for such reliance on modal closeness, one may claim that with the same (or even less) amount of change that we can make Lottie's belief undetectably false (and, thus, unsafe) we may be able to make the belief in many ordinary cases of knowledge undetectably false as well. In other words, in many ordinary cases of knowledge, if we try enough, we can think of *very close* possible worlds in which the belief is undetectably false (though, like Lottie case, such worlds may be relatively very few). For example, imagine someone calls her grandpa and after hanging up the phone forms the true belief that he is watching TV. This is an ordinary case of knowledge, yet we can think of very close possible worlds (though they may be rare) with a very small change from the actual world—arguably as small as changing the movement of several balls in the lottery machine to make Lottie's belief false—like a small change in the grandpa's blood circulation that leads to a stroke, a short-circuit somewhere in the electrical grid that leads to a blackout, etc., that can make the subject's belief undetectably false.

In response to a similar worry about Ernie case, Pritchard (2012) argues that this is not a problem for safety account since in case the details are set in a way that allows such modally close errors, we will lose the temptation to think the subject has knowledge in the first place. However, the claim here is not confined to a few cases that Pritchard easily may dismiss as not being genuine cases of knowledge, rather the point is that for *many* ordinary cases of knowledge, on reflection, there seem to be such modally very close errors (though they may be very improbable and, at first, hard to detect), and Pritchard response will have the unfavorable result of denying knowledge assignment for those cases.

II.2.3 Lack of Useful Predictions

In close relation with the previous problem, the reliance on the manifest image (and modal closeness) may also have to do with the inability of popular accounts like safety to make useful predictions in many cases. According to Williamson (2009), in order to decide whether a belief is safe, one should already have an idea of what knowledge is as well as whether it obtains. In other words, in deciding which close possible worlds are relevant and which changes are small we rely on our prior intuitions about the knowledge status of the belief. Thus, we do not usually get from safety to knowledge but vice versa which, in turn, makes accounts like safety unable to provide predictions in many cases where different considerations pull in divergent directions (Williamson, 2009).

II.2.4 Rejection of Knowledge in Statistical Cases

As I argued earlier, statistical evidence in some cases like Lottie is extremely reliable, far more reliable than many ordinary kinds of evidence that we base knowledge on. Nonetheless, many epistemologists, following their intuitions, have encouraged accounts of knowledge that exclude cases like Lottie and seriously criticized accounts like some versions of reliabilism for allowing the possibility of knowledge assignment for such cases. However, considering numerous failed attempts to provide a rationale—independent from the related intuitions—for our intuitive judgments about statistical cases, on the one hand, and the apparent reliability of such statistical evidence on the other, one may suspect that there is something wrong with the manifest image of knowledge. In that case, allowing the assignment of knowledge to cases like Lottie should, in fact, be seen as a feature for a proposed account of knowledge and not a bug.

II.2.5 Unreliability of the Manifest Image

As the discussion of the last chapter has shown, although the 2nd and 4th scenarios satisfy the criteria for accounts of knowledge like causal theory, sensitivity, and safety, some people may still find such knowledge attribution counterintuitive which is at odds with the claim of such accounts that they track the intuitions. Moreover, while the 2nd and 4th scenarios have a similar status with respect to safety (or sensitivity), they trigger very different intuitions in many people. In fact, people's intuitions may diverge even with respect to a single case (e.g. some have the intuition that Arash has knowledge in the 2nd scenario, while others may disagree).

Previous scenarios are only some examples of many cases in epistemology that trigger conflicting intuitions even among professional epistemologists. Accordingly, one may argue that some of the problems in epistemology arise from relying too heavily on such unreliable—and sometimes misleading—intuitions about knowledge.

In order to show how inquiries for a coherent analysis of knowledge can be misled by intuitive judgments, the following illustrates an example of how two cases can have a similar epistemic status but a different manifest image:

> Imagine Arash is in a park while knowing the following:
>
> (a) There are exactly 1 million animals in the park: 999,999 sheep and 1 dog in a random place among them;
> (b) All the animals in the park indistinguishably look like a sheep (i.e. the dog is carefully disguised as a sheep). And he has gained this knowledge by checking all the animals in the park one by one.
>
> **5th scenario:** Suppose Arash knows that one of these 1 million animals is behind a wall in front of him (e.g. he hears an animal-like sound, etc., and justifiably believes so). He forms the belief that there is a sheep there behind the wall. It is indeed a sheep.

6th scenario: Now imagine the animal moves from behind the wall so Arash can see that *as he already knew* it looks like a sheep. He forms the true belief that there is a sheep there.

Compared with the 5th scenario, far more people have the intuition that Arash has knowledge in the 6th scenario (which is, in fact, the equivalent of the 2nd scenario). Nevertheless, in the 6th scenario, Arash does not seem to be in a better epistemic position. This is because perceiving the animal—once again—in the 6th scenario, does not provide him with anything that he didn't already know and therefore does not seem to give any better grounding for knowledge.

It may be tempting to say that the 5th and 6th scenarios are epistemically different because the 6th scenario includes a causal connection that the 5th scenario lacks. However, even in the 5th scenario, Arash's belief is causally connected with the facts: The sheep behind the wall has already caused him—when he was checking the animals—to have a perception similar to the 6th scenario. The following argument shows whoever assigns knowledge to Arash in the 6th scenario, is committed to assigning knowledge to him in the 5th scenario as well, despite it being counter-intuitive:

P1: There are exactly 1 million animals in the park
P2: There is an animal behind the wall in the park
C1 (from P1 & P2): One of these 1 million animals is behind the wall
P3: Arash has already perceived all 1 million animals
C2 (from C1 & P3): Arash has already perceived the animal behind the wall
P4: Arash, in this park, has the knowledge that an animal is a sheep if his true belief is based on perceiving an animal as sheep (consistent with the knowledge assignment in the 6th scenario)
C3 (from C2 & P4): Arash (in the 5th scenario) *knows* that the animal behind the wall is a sheep

The 4th premise of this argument should be accepted by anyone who thinks Arash has knowledge in the 6th (or 2nd) scenario. Also, the first three premises are based on perception and memory which according to Goldman (1967) are "appropriate" causal processes. However, one may object that what is demanded is not just some sort of causal connection, but a specific one, namely between the belief on the one hand, and *the very fact believed* on the other. She may argue that it is true that in the 5^{th} scenario the sheep behind the wall has been ultimately (causally) responsible for Arash's perceptual belief as well as his belief that one of those animals is behind the wall, yet the perceptual belief in the 5^{th} scenario has happened before the sheep is selected so it is not the case that the very fact that "there is a sheep behind the wall" caused him to believe there is a sheep there.

The problem with this objection is that it leads to denying knowledge assignment for ordinary cases like the following: Imagine a similar scenario as the 5^{th}, with the small difference that now all the animals in the park are real sheep (and Arash knows it). Now, if the objection is true, we should deny that he knows there is a sheep behind the wall even in this new scenario, but the knowledge, in this case, seems an ordinary case of knowledge.

In another attempt, one may argue that the 5^{th} scenario is epistemically inferior to the 6^{th} scenario because unlike the latter which includes a direct and immediate perception, the former involves a long deductive inference which many may fail to make. In reply, this does not show that Arash does not have knowledge in the 5^{th} scenario (or that 5^{th} scenario is epistemically inferior to 6^{th}), instead, it explains why we may (mistakenly) think so. In any case, it seems that even after we are presented with the argument, the intuitive difference between two scenarios does not completely go away.

The last attempt to epistemically distinguish 5^{th} and 6^{th} scenarios is to claim that the belief in the 5^{th} scenario, unlike the 6^{th}, is not sensitive and/or safe. This attempt fails too since given that these two scenarios are identical in every aspect except for the presence of a wall in the 5^{th} and considering that in both cases there are some causal connections between the facts and the belief, if one accepts that Arash's belief in the 6^{th} scenario is sensitive and/or safe, it seems very hard to deny it for the 5^{th} scenario. This is because any amount of change that can turn Arash's belief undetectably false in one of the scenarios, can make it false in the other as well; if the change from the actual circumstance is the removal of the sheep, in both scenarios Arash would not form the belief anymore, and if the change is to replace the sheep with a disguised dog, in both cases he would still form the false belief that there is a sheep there.

II.3 MAKING SENSE OF THE MANIFEST IMAGE

As noted earlier, the 2nd, 4th, 5th, and 6th scenarios have a similar epistemic status but trigger different intuitions. It seems that the intuitive difference between these cases is due to the presence of direct and immediate perception—as a natural, familiar and ordinary way of forming a belief—in the 2nd and 6th scenarios (and its absence in the 4th and 5th scenarios). After all, it is very natural to form the belief that there is a sheep there when one directly perceives an animal that looks like a sheep.

Accordingly, we feel more comfortable to assign knowledge to the 2nd and 6th scenarios in which the evidence includes a direct and immediate perception (albeit such perception, considering that the dog is disguised as a sheep, is not helpful at all) because such direct perception brings a strong intuitive force to the table which in turn can eclipse the "unfavored" statistical aspect of the evidence.

On the other hand, we are less comfortable to assign knowledge to the 5th and especially 4th scenarios since in those cases the statistical part of the evidence is more dominant and takes up more attention due to the absence of direct and immediate perception and its resulting intuitive force in support of the belief.

Also, the divergence in intuitions about each case can be explained in reference to the fact that the strength of the intuitive force resulted from the immediate perception, on the one hand, and the salience of the statistical part, on the other, may vary among different people and in different contexts.

Such comparisons between cases which are epistemically similar but intuitively different illustrate how the manifest image of knowledge may be biased in favor of ordinary and non-statistical evidence. In what follows, I discuss some possible sources of such bias against statistical evidence in our intuitions.

II.3.1 Historical Absence of Statistics

To better understand the manifest image of knowledge—and biases that come with it—we should consider that, as Sellars explains, both manifest image and scientific image come with a history; the main outline of the manifest image has shaped in pre-history while the scientific image has formed before our eyes (Sellars, 1963).

Accordingly, considering that the concept of probability, as we know it, and statistics are relatively recent[3] without which we couldn't make sense of probabilistic systems or produce statistical evidence, I think the bias has to do with the fact that we didn't have access to (explicit) statistical evidence during most of our history when our intuitions and the common-sense notion of knowledge were formed. Compare this with ancient, familiar and ubiquitous sources of knowledge, like perception, testimony, memory, and even induction, that have been used from pre-history to understand and make predictions about the world. Even today, we don't deal, in our daily lives, with statistical evidence as much as we deal with other kinds of evidence. Arguably, if statistical evidence has been somehow accessible and prevalent from pre-history among our ancestors, it may have been incorporated in the 'meaning' and the manifest image of knowledge.

[3] Although historical records show that during much of our history we have used die-shaped animal bones and other randomizing objects to gamble for entertainment, communicating with gods or predicting future, the mathematical theory of probability famously started in the mid-17th century when a gambler's question led to a correspondence between Blaise Pascal and Pierre de Fermat. See: Hacking, 2006.

II.3.2 Poor Probabilistic Judgments

One may also argue that this bias is partly the cause and/or symptom of the fact that we have a poor performance in making probabilistic judgments and are prone to many biases; in addition to famous cases like the Monty Hall problem, Gambler's fallacy, etc., some studies show that judgments mediated by intuitive heuristics make egregious mistakes even in most simple and basic qualitative laws of probability. For instance, although the conjunction rule states that the probability of a conjunction, P(A&B), cannot be greater than the probabilities of its constituents, P(A) and P(B), the availability and representativeness heuristics can make a conjunction intuitively appear to be more probable than one of its constituents (see: Tversky and Kahneman, 1983).

II.3.3 The Salience of Uncertainty vs. Impression of Certainty

One may argue that the presence of the wall in the 4th and 5th scenarios—apart from depriving Arash of the intuitive force of ordinary ways of forming belief like perception—may lead to some negative feelings of doubts and make the possibility of error and uncertainty more salient. Accordingly, the bias may partly have to do with the salience of error and uncertainty in statistical cases compared with the psychological feeling of certainty associated with common kinds of evidence. As an interesting example, in a study, one-third of a sample of judges who have been asked to translate the famous *beyond reasonable doubt* standard into numerical probabilities, translated it as a probability of one.[4] This shows how some non-statistical evidence—such as eyewitness—can produce an impression of certainty despite not being much reliable[5]. Although many of us, when asked, would admit that there is some possibility of error associated with many kinds of evidence, this possibility does not usually

[4] Thomson (1986) mentions this from Simon (1969)

[5] In a famous case, of all the US prisoners who have been exonerated based on DNA evidence, 70% of them were convicted based on mistaken eyewitnesses making eyewitness misidentification the greatest contributing factor to wrongful convictions. See: Innocence Project, 2017.

take up as much attention as in the statistical cases in which even the most improbable errors are very salient and explicit.

II.3.4 The Impact of Cognitive Biases

The salient possibility of error in statistical cases may get further focus or magnification due to the impact of some cognitive biases. For instance, with respect to cases like Lottie, there is a range of cognitive biases that can lead to overestimation of Lottie's infinitesimal—though salient—chance of winning (i.e. the chance of forming a false belief). On the one hand, our cognitive limitations regarding very small (or very big) numbers prevent us from truly grasp how small the chance of winning is. On the other hand, the fact that people usually tend to assign a higher chance of success to themselves than average amount and biases like optimism bias, choice-supportive bias, self-serving bias, neglect of probability bias, etc., lead to overestimation of the chance of winning; and lottery slogans like: "It could be you", "Just Imagine!", "Luck is right in front of you, play.", etc. are not much of a help!

The impact of such biases may become clearer by considering a modified version of Lottie case in which Lottie has bought all the tickets except one ticket and forms the belief that she has a winning ticket. Many people would find this modified version intuitively different and intuitions about it may be less harsh. Some may even think she has knowledge in this case (though, even in this case, some cognitive biases can magnify the infinitesimal chance of losing).

II.4 Moving Towards A Scientific Image of Knowledge

As Sellars (1963) explains, from the point of view of scientific image, the manifest image is "inadequate" though is still pragmatically useful and provides the initial framework that a scientific image can grow out of. In what follows, I turn back to the safety account of knowledge as an example and briefly discuss how some possible moves towards a more scientific image can help with some of the aforementioned issues.

In explaining the criteria for safety Pritchard suggests:

> In wanting our cognitive success to be immune to luck we are not thereby desiring that it be free from any possibility of error, no matter how remote. Accordingly, as the error becomes more remote—that is, as more needs to change about actual circumstances for the agent to (counterfactually) form a false belief—so we become more tolerant of it, to the point where we no longer regard the counterfactual error as indicating that there was anything lucky about the target cognitive success. The anti-luck intuition thus manifests itself, in keeping with how we are reading the safety principle, with a complete intolerance of error in close counterfactual circumstances, a tolerance of error in remote counterfactual circumstances, and a sliding scale of tolerance between these two extremes (2012, p. 255).

Interestingly, Pritchard (2007) initially suggests and defends a different version of safety as a condition that captures the anti-luck intuition; that version of safety lacks such extreme of *complete* intolerance of error in close counterfactual circumstances and only requires that the belief be true in *most* near-by possible worlds. However, later in that paper, he adds such extreme to the characterization of safety with the specific goal of accounting for the intuition that cases like Lottie are not cases of knowledge (Pritchard, 2007).

That said, one may claim that the intuition against cases like Lottie is not, in fact, an anti-luck intuition rather the result of other factors like cognitive biases, etc., and, thus, we do not need to account for it by an anti-luck condition such as safety in the first place. Putting that aside, given the reliability of the belief in cases like Lottie as well as the fact that there is no good rationale for denying knowledge assignment for such cases, positing such extreme—which in turn excludes knowledge assignment in many statistical cases like Lottie—does not seem well justified. Instead, it seems more natural to drop this extreme and go back to the initial version of safety or, alternatively, posit a continuing sliding scale according to which the closer the error is, the *less* it is tolerated (i.e. tolerance tends towards zero but never reaches it).

Such strategies will have the benefit of allowing knowledge assignment for statistical cases like Lottie but still appeal to the modal closeness of the error which does not seem to be the best way of measuring the amount of luck, risk or error, considering that we have access to better alternatives like probability theory.

That being so, one may move even further towards a scientific image by going beyond our intuitive tendency to care about the modal closeness of error and construing “could not easily” clause in the criteria for safety as referring to a low probability of error rather than modal closeness of it. Although defining safety in this way seems very plausible and natural, I think due to its counter-intuitiveness it is not taken very seriously.

An account of knowledge that goes beyond intuitions in this way, in addition to allowing knowledge assignment in many statistical cases, provides extra benefits such as making better predictions: Although finding the probability of error itself is not always straightforward, compared with the reliance on modal closeness in the original safety account, it is more independent from prior intuitions about knowledge status of the belief and thus less circular. Additionally, it leads to similar treatment of Lottie and Gottie cases which seems the right approach.

II.5 WHY NOT ABANDON THE IDEA OF KNOWLEDGE ALTOGETHER?

One natural reaction towards the aforementioned problems with the manifest image of knowledge would be to doubt the coherence of the notion of knowledge and opt for abandoning it altogether. David Papineau, for example, argues that a concern with knowledge, rather than true belief, is “a stone-age hangover” which does “an appreciable amount of harm”; he thus asks us “to stop thinking in terms of knowledge” (Papineau, forthcoming, pp.2-3).

In response, although abandoning the notion of knowledge is on the table, considering the central role of knowledge in our thought, language, and culture, it should be seen as the last option. A better reaction, I suppose, is to give up the problematic intuitions involved in the notion of knowledge but still keep the other unproblematic ones.

I agree with Pritchard that our thinking about knowledge is governed by two master intuitions that distinguish knowledge from mere true belief. Pritchard calls the first one *anti-luck intuition* according to which when one has knowledge, her believing truly (i.e. cognitive success) is not a matter of luck. According to the second intuition, which Pritchard calls *ability intuition,* one has knowledge when his cognitive success is the product of his relevant cognitive abilities (Pritchard, 2012).

I also agree with Pritchard when he argues that although these two intuitions seem to be two "faces" of a single intuition, we need to conceive of them as "imposing distinct epistemic demands, and hence as requiring independent epistemic conditions" (Pritchard, 2012, p. 249). Such independence between these two intuitions enables us to treat them differently. Accordingly, considering that none of the aforementioned problems with the manifest image of knowledge seems to have its root in the ability intuition, we can keep this seemingly unproblematic intuition in our thinking regarding what distinguishes knowledge from true belief.

However, as noted above, the anti-luck intuition (especially its reliance on modal closeness which underlies the original safety condition) seems to be behind many of the aforementioned problems with the manifest notion of knowledge. Nonetheless, by decreasing the commitments to some of the underlying intuitions and characterizing safety in a more independent way (say, in reference to the probability of error rather than modal closeness), we may be able to keep some version of anti-luck condition as well.

It is worth noting that in order to defend the role of these two independent master intuitions in our thinking of knowledge, Pritchard refers to a popular story about the genealogy of the concept of knowledge according to which the notion of knowledge has evolved as a practically useful mean to help us discover truth on matters that interest us. According to this picture, for a true belief to be considered as knowledge, first, it should be the product of the agent's relevant cognitive abilities and, second, its truth should not be substantively due to luck (2012, pp. 275-278). That being so, matters like allowing knowledge assignment for cases like Lottie (which satisfies both requirements: The belief is the product of Lottie's relevant cognitive abilities and its truth is not *substantively* due to luck) as well as replacing the reliance on modal closeness with a better alternative not only can be consistent with such historical picture but also can help the notion of knowledge in its original goal of providing practical benefits.

Moreover, on reflection, there is a chance that the counter-intuitiveness associated with such modifications in the notion of knowledge gradually go away. This is because, as I argued in previous sections, one main reason behind our intuitive bias against statistical evidence is that, historically speaking, we haven't had much experience of dealing with this kind of evidence and, thus, the gradual increase in our exposure to statistical evidence may lead to internalization of relying on it and a change in related intuitions which can be seen as another step in evolution of the concept of knowledge.

There are many other examples of how our prior intuitions have updated in the light of new discoveries. For instance, for a long time, it was very intuitive that the sun is circulating around the earth and claiming the opposite would seem very counter-intuitive; however, nowadays we don't find the claim that earth circulates around the sun much counter-intuitive.

In short, an updated version of anti-luck condition (which may even gradually gain strong intuitive support) as well as the original ability condition can plausibly be considered as what makes knowledge more than mere true belief. Now, considering such possibility of an unproblematic notion of knowledge, on the one hand, and the central role of knowledge in our lives, on the other, abandoning the notion of knowledge does not seem to be much defensible.

II.6 CONCLUSION

Those of us who are not fans of abandoning the concept of knowledge, or possibility of an analysis of it, are left with two options. The first one is to hope that someday we will see a satisfactory analysis that fully conforms to the manifest image of knowledge. However, on the one hand, this option does not seem very promising given the long history of failed attempts; on the other hand, it does not seem to be very helpful in resolving many problems that seem to be rooted in the manifest image of knowledge itself.

In this paper, I suggested a second option: widening our search area to include accounts of knowledge that may not be fully compatible with our (current) intuitions. I recommended some historical considerations to better understand the manifest image and, briefly, explored the possibility of a move towards a more scientific notion of knowledge.

Obviously, there are numerous considerations to be taken into account in proposing any account of knowledge (Gettier cases, pragmatic encroachment, skepticism, etc.). Here I had neither the aim nor the space to provide a full-fledged scientific image of knowledge; my goal, instead, was to provide a starting point for exploring the possibilities that arise from decreasing commitments to manifest image. Possibilities that allow accounts of knowledge that are more rational, coherent, and make better predictions. In the next chapter, I will turn to a parallel issue with statistical evidence in the context of legal cases.

III. Chapter 3: Statistical Evidence in Courtrooms

III.1 Introduction

There is a close relationship between the discussion of statistical evidence in epistemology and in the context of legal cases. On the one hand, it seems that the same counter-intuitiveness behind denying knowledge assignment for statistical cases in epistemology is playing an integral role in making us uncomfortable to base conviction on statistical evidence in courtrooms. On the other hand, many proposals to discriminate between statistical and individualized evidence in legal cases, stem from familiar ideas in epistemology such as causal connection, sensitivity, safety, etc.

That said, the issue of statistical evidence in legal cases is not merely an epistemological matter, and there are various social and legal considerations in play—from concerns about the deterrent effect of law to legitimate worries about racial profiling, etc. However, in this thesis, putting those factors aside, I merely focus on whether statistical evidence can provide an acceptable basis for conviction.

There are a number of *standards of proof* that are referred to in courts. The relevant standard in civil trials is the *preponderance of evidence* which is met when the claim is more likely true than false (typically quantified as the likelihood of more than 50%). A more demanding standard employed in criminal trials famously requires that guilt of the defendant be established *beyond reasonable doubt* (usually quantified as 90-95% likelihood) (Gardiner, forthcoming).

The following is an example of a possible civil case:

Smith v. Red Cab

Mrs. Smith was driving home late one night. A taxi came towards her, weaving wildly from side to side across the road. She had to swerve to avoid it; her swerve took her into a parked car; in the crash, she suffered two broken legs. Mrs. Smith, therefore, sued Red Cab Company. Her evidence

> is as follows: she could see that it was a cab which caused her accident by weaving wildly across the road, and there are only two cab companies in town, Red Cab (all of whose cabs are red) and Green Cab (all of whose cabs are green), and of the cabs in town that night, six out of ten were operated by Red Cab. (Thomson, 1986, p. 199)

Although, in this case, the likelihood of the Red Cab being the cause of the accident seems to exceed the required 50% for civil cases (I'll express reservations about this claim in the next chapter, but let's assume it is true for now), such statistical evidence is not usually an acceptable basis for conviction in courts, and many people are uncomfortable with basing conviction on such evidence.

In this chapter, I discuss two attempts to discriminate between individualized and statistical evidence by Judith Thomson and Martin Smith. I will argue that while such proposals may track the way the intuitions work, they fail to provide a rationale, independent from intuitions, for preferring individualized evidence over statistical evidence. I will, then, argue against the claim that the probabilistic content of statistical evidence makes it inferior to individualized evidence.

III.2 THOMSON: CAUSAL CONNECTION

Thomson discriminates between individualized and statistical evidence by claiming that the former, unlike the latter, is causally related, in an appropriate way, to the fact it is presented to support. She explains that such individualized evidence is either "backward-looking" when the fact causes the evidence, for example, in case the accident causes the eyewitness testimony; or "forward-looking" when the evidence causes the fact like when drunk driving causes the accident; or there is a common cause involved like when a drunk driver causes the accident and a later accident (Thomson, 1986).

No doubt that we are more comfortable with cases in which there seem to be a causal connection between the fact and the evidence (or the belief, recall Goldman's causal theory of knowledge), however, as I discussed in chapter 2, this phenomenon can be explained in terms of various factors like by considering that seeking casual structures—which is especially very efficient in learning based on limited input[6]—has been a familiar way of understanding and controlling the environment for most of our history.

Such historical considerations that can explain a possible bias against statistical evidence in intuitions, together with the apparent reliability of such evidence cast doubt on the reliability of the intuitive judgments against statistical evidence. That being so, the opponents of statistical evidence are expected to provide a rationale, independent from the intuitions, for preferring the individualized over the statistical evidence.

Thomson seems to provide such a rationale. She claims that in the case of individualized evidence, we take ourselves to have a guarantee of truth (though she points out that such evidence is not a deductively valid proof), while statistical evidence leaves room for luck in an unacceptable way (Thomson, 1986).

However, it is not completely clear what she exactly means by a guarantee of truth. The view that there can be a guarantee of truth only if there is *indeed* a causal connection between the evidence and the fact is not much of a help here. This is because although such view can be seen as a respectable externalist theory of epistemic justification (according to which whether a belief is justified or not is partly determined by external factors like whether the evidence is indeed causally connected with the fact), it will not help us to distinguish the individualized evidence and statistical evidence in real situations (like legal cases) since in such cases we don't have access to those external factors[7].

[6] For related computational models see: Tenenbaum et al., 2011. See also: Orban et al., 2008.

[7] Gardiner (forthcoming) makes a similar point. Also, a similar line of argument can be used against the view that what distinguishes the individualized and statistical evidence *in legal cases* is that the belief in the former, unlike the latter, is sensitive. More on this view: Enoch et al., 2012; to see the argument against it: Smith, 2018; Gardiner, forthcoming.

In other words, in real cases of individualized evidence, the evidence itself does not suffice to decide whether there is indeed a causal connection between the fact and the evidence without which we can't take ourselves to have a guarantee of truth. For example, imagine that there is an eyewitness who claims that he saw Joe stealing from the store. Now it does not follow from the evidence (i.e. the eyewitness testimony) that it is indeed the fact that Joe was stealing from the store that caused the eyewitness testimony since there are other possibilities like misidentification, hallucination or even lying; thus we cannot take ourselves to have a guarantee of truth on the basis of such individualized evidence (in fact, if the court was in a position to know whether Joe's stealing indeed has caused the evidence, it didn't need the testimony of the eyewitness in the first place!).

It is also worth noting that, as J.L. Mackie argues, partial causes are neither sufficient nor necessary for the effect; rather, following J.S. Mill, only the whole state of the universe (or at least the whole spatiotemporal area that can possibly have an effect) prior to an effect can be sufficient for it (Mackie, 1965; Crane, 1995). It follows that—unless one is omniscient—causation cannot and should not give a guarantee of truth.

For example, Thomson explains that if it turns out that on the evening of the accident the Red Cab company had held a party for its drivers that led to a drunken brawl, this would provide an individualized evidence against Red Cab company because the party (as the evidence) can causally explain the fact that Red Cab caused the accident (Thomson, 1986, p. 203). However, on the one hand, the fact that Red Cab is responsible for the accident could obtain without there being any party (e.g. an individual driver from Red Cab could be part of a drunken brawl somewhere else); thus, the party is not a necessary cause behind the fact. On the other hand, there could be the party and the brawl without the fact being obtained, namely, without the Red Cab being the cause of the accident; hence, the party is not a sufficient cause behind the fact.

Accordingly, individualized evidence does not seem to provide any special guarantee of truth. Although relying on causal explanations in the case of individualized evidence can be a reliable method of increasing the chances of drawing the right conclusions, this is the case for the statistical evidence as well. What seems to distinguish them though is the fact that the former, unlike the latter, is associated with a psychological feeling of certainty, trust, and comfortableness which I talked about in chapter 2.

Consequently, Thomson's claim that in the case of individualized evidence we take ourselves to have a guarantee of truth seems to be at best a good description of our psychology but does not provide an independent rationale for preferring individualized over statistical evidence. For if we take this psychological feeling of certainty away, what would be left of that guarantee of truth? This probably has to do with the fact that the more conscientious we are in drawing a conclusion, by elaborating different possibilities, the less certain we will end up being.

III.3 SMITH: NORMIC SUPPORT

Smith suggests that individualized evidence, unlike statistical evidence, provides normic support. He explains that "a body of evidence E *normically* supports a proposition P just in case the circumstance in which E is true and P is false would be less normal, in the sense of requiring more explanation, than the circumstance in which E and P are both true" (Smith, 2018, p.1208).

For example, imagine there is an eyewitness testimony that a taxi from Red Cab is behind the accident; now, given this testimonial evidence, in case the taxi is from the Green Cab (i.e. the conclusion is false), we will look for more explanations compared with the case when it is from the Red Cab. Whereas when the evidence is statistical, we don't seem to look for a special explanation in case the taxi is from Green Cab.

Maybe as Gardiner (forthcoming) suggests, the reference to normic support can shed light on the sense of guarantee that we seem to associate with individualized evidence. Admittedly, we have a preference for evidence that provides normic support, and we may even feel a sense of guarantee or certainty with respect to such evidence. However, this may be merely a psychological feeling and the difference between individualized and statistical evidence may be the result of our different psychological attitudes towards these two kinds of evidence: Due to a psychological feeling of certainty in the case of individualized evidence, if the result is different from what is expected, we become surprised and look for further explanations; whereas it is not the case when the evidence is statistical because not only statistical evidence lacks the associated psychological feeling of certainty but also saliently reveals all the available possibilities no matter how much improbable they are.

Additionally, Smith's view can be seen as putting emphasis on one of the factors that has historically shaped our intuitions, namely the explanatory and predictive value of data. In ordinary cases when the result is different from our expectation, even if we are told about different possibilities, we may still tend to look for further explanations because it has many practical benefits such as improving our future predictions, while it is not the case for many statistical cases like Lottie in which when the result is different from the expectation (i.e. when Lottie wins), finding explanations—say, by figuring out what physical processes happened in the lottery machine that led to her number instead of a different one—is not an efficient and reliable way of improving future predictions about winning numbers[8].

That said, our judgments about what counts as *normal* (or abnormal and, thus, in need of explanation) seem to be affected by many factors (like our past experiences, common

[8] It is worth noting that even in some cases of statistical evidence, looking for further explanations when the expectations do not fulfill has some practical benefits (like revising the models, updating prior probabilities, etc.) and help with improving the probabilistic predictions.

sense expectations, etc.) some of which may be unreliable. In fact, Smith emphasizes that this notion of normalcy is not the statistical notion of what frequently obtains; he also alludes to the possibility for a body of evidence to provide normic support without providing much probabilistic support (Smith, 2018). Accordingly, one may argue that such intuitive notion of normalcy associated with individualized evidence, in fact, puts statistical evidence in a better position than individualized evidence and not vice versa. In any case, although Smith's account offers an interesting description of our intuitions and, to some extent, an interpretation of the legal practice, it does not seem to provide an independent rationale for privileging individualized over statistical evidence.

III.4 PROBABILISTIC CONTENT

One may argue that the reason why statistical evidence is inferior to individualized evidence (or even the reason why it does not provide normic support) is that it adds an extra layer of uncertainty, namely it has a probabilistic content. For example, the content of Lottie's belief is like "my ticket is *almost certainly* a loser", or in *Smith v. Red Cab* case the content of the judgment is like "*with the probability of 0.6* the cab which caused the accident was operated by Red Cab", while in the case of an individualized evidence such as an eyewitness testimony the content seems to lack the probabilistic element and is like: "the cab which caused the accident was operated by Red Cab".

In response, the mere fact that the content of the judgment in statistical cases is probabilistic does not make statistical cases, in general, less certain or reliable than ordinary cases (compare a judgment based on a not much reliable eyewitness testimony with a judgment based on statistical evidence which has an extremely high probability). In fact, if the amount of uncertainty in the content is minimal (i.e. the probability is very high), it would have a negligible impact on the overall uncertainty pertaining to the judgment.

Besides, even in non-statistical cases, the strength of the belief and our confidence in the evidence are sometimes reflected in the content of the judgment adding a (usually qualitative) probabilistic aspect to it—which has benefits like more communicability of the involved uncertainty. For example, when we don't have enough information, or rely on, say, a not completely reliable news agency, or an old memory, etc., we may communicate this unsureness through the content by adding to it phrases like "it is very probable…", etc. On the other hand, the probabilistic aspect of the content in statistical cases could alternatively be reflected in, say, the strength of the belief (e.g. in Lottie case, she could alternatively form a *strong belief* with a non-probabilistic content like "the ticket is a loser"). Accordingly, a reference to the probabilistic content of statistical cases does not seem to provide a rationale for discriminating against statistical evidence.

III.5 CONCLUSION

In the previous chapter, I discussed some problems that arise from relying too heavily on the intuitive picture of knowledge. The same is true in the context of legal cases. The strong counter-intuitiveness of basing judgments on statistical evidence may put attempts to discriminate against statistical evidence in danger of leaning towards some sort of post-hoc reasoning to justify and support our already made intuitive judgments about what can be counted as an acceptable basis for conviction, rather than an unbiased evaluation of different consideration is each side.

In this chapter, I argued that Thomson's emphasis on causal connection and Smith's point about normic support can at best track the way our intuitions react differently to statistical and non-statistical cases but does not seem to provide a rationale for discriminating against statistical evidence. I then explained how differentiating between the two types of evidence in reference to the probabilistic content of statistical evidence fails.

That said, producing reliable judgments based on statistical evidence requires caution and enough confidence that the criteria required for a plausible application of statistical methods are satisfied. In the next chapter, I will discuss this issue along with some worries about the reliability of statistical methods.

IV. Chapter 4: On Reliability and Applicability of Statistical Methods

IV.1 Introduction

Even if indeed there is no rationale for rejecting statistical evidence as a basis for knowledge or conviction, there are still legitimate concerns about that. One worry has to do with the reliability of the statistical methods that produce the relevant statistical evidence. One may ask how do we know the current statistical methods are reliable and result in accordance with what they claim they do—not only on the paper but in concrete situations? The second concern pertains to the applicability of such methods. According to this worry, even if we grant that statistical methods can be applied in appropriate situations, how can we decide in which cases we can apply them? In the next two sections, I will address these two concerns in turn.

IV.2 Reliability of the Statistical Methods

Before turning to the case of statistical methods, it is useful to consider how we respond to other reliability questions, say, regarding the reliability of our visual perception. One obvious answer to such questions is in reference to our experiences that show they have worked reliably. If one finds out that her visual system does not work properly (i.e. it produces misleading results), she will stop considering her eyesight as a reliable faculty.

However, in order to give a verdict on the reliability of a specific method, we need more systematic approaches (personal and unsystematic approaches themselves may not be always very reliable which explains why some people take palm reading, etc., as reliable methods of acquiring true belief!). Science is one of the most systematic practices that we have ever accessed. Now not only the reliability of statistical methods and their applicability to various concrete situations have long been tested by science, but also probabilistic language has become pretty much the language of science through which scientists communicate the results of different studies.

On top of that, even in our daily lives—from the weather forecast to the stock market—we are constantly relying on the results produced by statistical methods. In one sense, such statistical methods—and the resulting statistical evidence—have even more applicability and are more fundamental than many ordinary kinds of evidence. For one thing, scientists use such methods along with other mathematical machinery in areas that are beyond the scope of many ordinary methods of gaining knowledge (e.g. in some theoretical areas of physics); for another, the study of the reliability of various methods—including ordinary ways of gaining knowledge like eyewitness, memory, etc.—itself hinges on statistical methods and the results of such studies are produced and communicated in terms of statistical information (e.g. eyewitness is n% reliable).

In response, one may grant that we have very strong evidence from various sources that statistical methods can be reliably applied, outside the math class, to many concrete cases but still express the worry that we don't have the same confidence in statistical methods which are relatively recent compared with other kinds of evidence which we have relied on from prehistory. However, as I argued in chapter 2, such historical considerations partly explain why we find the reliance on statistical evidence counter-intuitive and support the claim that we are biased against statistical evidence. That being so, now that we do have access to such evidence and have good reasons to think it leads to reliable results, we are justified to base judgment on.

That said, the claim that the roots of statistical inferences in our history and cognition are not as deep as other kinds of evidence is not tenable. In fact, considering that most of the cognition is automatic and outside of consciousness (Bargh and Chartrand, 1999), there is no surprise that we are usually unaware of statistical processes that underlie our cognition. For example, it is widely accepted among various theories of learning that in order to acquire generic knowledge, we need to make generalizations from newly perceived information—involving a specific time and location, and particular objects—to new situations, objects, etc., and that these generalizations rely on statistical procedures involving sampling several episodes of experience to find the relevant information that can be generalized (Csibra and Gergely, 2009).

More specifically, some studies suggest that Bayesian norms underlie many of our general cognitive capacities as well as many specific ones, that result from rapid, reliable and unconscious processing, including language, perception, memory, etc. (although it is not the case for tasks that require explicitly and consciously manipulating of numerical probabilities) (Tenenbaum et al., 2011).

This, if true, has a crucial implication, namely, there is no difference, in principle, between individualized evidence and statistical evidence, and what we take to be a historical and ordinary kind of evidence may actually be statistical information in disguise. What seems to distinguish them, though, is that in the case of ordinary evidence we are not aware of the underlying statistical aspects which, in turn, may result in an impression of certainly as well as more confidence and comfortableness; whereas the awareness about the statistical aspects in the case of statistical evidence affects our psychology and leads to uncomfortableness, distrust, and counter-intuitiveness.

In a brief digression, one may wonder if such statistical aspects really underlie many parts of our cognition, why aren't we able to consciously access or communicate them? In response, I think this can be understood in reference to the adaptive value of such conscious access and the pressures for communicative efficiency which usually result in conveying the information as concisely as possible (for example, Zipf's law states that the length of words in a language is mainly determined by their frequency of use[9]). Accordingly, one may argue that given the scarcity of cognitive resources, it has not been optimal to have conscious access to the underlying statistical and probabilistic aspects of the loads of information that we constantly acquire from the environment. Also, considering most of the cases in most of our history, conscious access to *explicit* statistical data, does not seem to sever much practical value. For example, it usually doesn't make sense that after seeing a sheep one explicitly think and communicate as "it is 99% probable that there is a sheep there" when simply saying "there is a sheep there" does all the job.

IV.3 APPLICABILITY OF STATISTICAL METHODS TO PARTICULAR CASES

Even if statistical methods are reliable, we are still in need of some criteria to decide when and where we can apply them. Before making any judgment based on the available statistical evidence, we need to be confident that the judgment is true or, more precisely, it has a low overall margin of error. In what follows I analyze this overall margin of error in terms of three layers of error in a way that error in lower layers has more significance than higher layers, to the extent that a high level of error in the lower layers can make the evidence irreverent, even if the level of error associated with higher layers is infinitesimal.

[9] A recent study of 10 languages suggests that the average information content predicts the word length better than frequency, but this can be considered as an improvement over Zipf's law. See: Piantadosi et al., 2011.

IV.3.1 Three Layers of Error

Apart from the statistical part, the body of evidence includes other information (coming from perception, memory, testimony, etc.) that are usually used as a basis for making further (statistical or non-statistical) inferences. The first layer (type) of error concerns with the chance of error in this background information (this error is associated with both non-statistical and statistical cases). For example, in order to make any statistical inference about losing/winning in a lottery, the body of evidence should include non-statistical data such as the fact that the subject has a lottery ticket, etc., and there is some chance of error in such data.

The second layer (type) of error pertains to how much the body of evidence includes main factors that are relevant in obtaining the fact. The more the body of evidence lacks some of the main relevant factors, the higher the margin of error would be sometimes to the extent that we should probably suspend any judgment (this type of error is also associated with both non-statistical and statistical cases). For example, imagine someone buys almost all the tickets in a lottery, but then she finds out that the lottery is not fair or someone is going to cheat. Apparently, in this situation, it does not make sense for her anymore to rely on the statistical evidence, namely, merely reflect on the odds involved and conclude that she will win; this is because in this case the statistical evidence is not sufficiently relevant (in fact, in case of cheating it would be completely irrelevant) and the body of evidence lacks some main relevant factors, and, thus, any judgment based on that will have a high margin of the second type of error.

Even if the statistical evidence is sufficiently relevant and plausibly applicable to a particular case, the statistical evidence (often explicitly) specifies a third layer (type) of error, that is, the probability of the judgment being false according to the evidence (only statistical cases deal with this type of error). For example, in Lottie case, given there are one million tickets overall and only one winning ticket, the statistical evidence predicts that the probability of error, namely the probability of Lottie forming a false belief, is one in one million. This kind of error is the one which usually gets the most attention because it pertains to what the available evidence predicts rather than what is missing from the body of evidence.

Accordingly, before basing any judgment on the statistical evidence in a particular case, we need to make sure that the overall margin of error is sufficiently low. Any error in the first layer, that is the error associated with the background information, can significantly affect the overall margin of error (e.g. if the background data that one holds most of the tickets in a lottery is erroneous). Similarly, if the second type of error is high (i.e. the evidence is not sufficiently relevant), we are not justified in using the evidence as a basis for the target judgment, even though the third type of error is extremely low.

IV.3.2 Some Criteria for Keeping the Second Type of Error Low in Case of Probabilistic Judgments

In case of a probabilistic judgment, in order to keep the second type of error low, firstly, we need to be confident (or we need to *know* if you like) that the fact that the evidence is about is obtained randomly, otherwise, the statistical evidence would not be relevant in the first place. Secondly, in case the evidence is the probability of specific items in a population, we can take the probability to be the proportion of those items in the population only if we are confident that each member of the population has the same chance of being chosen.

The following two cases illustrate when these criteria are or are not met:

> Urnie is presented with an urn in which there are 100 balls, 90 of them are white and the other 10 are black. He puts his hand inside the urn and picks one of the balls, but you don't know the result. You want to make a probabilistic judgment about the color of the ball that he picked based on the evidence about the number of black and white balls in the urn in the two following scenarios:
>
> **Fair Urn:** You have very good reasons to think that the balls inside the urn are distributed and selected randomly (the first criterion is satisfied), and you are confident that there is no

difference between balls that can affect their chances of being chosen, that is, each ball has an equal chance of being chosen (the second criterion is met).

Unknown Urn: You don't know whether the balls are distributed and selected randomly; for example. there is a good chance that Urnie has intentionally picked a black ball (the first criterion is not met). Putting that aside, you don't also know whether balls have a uniform chance of being chosen since you don't know whether the balls are similar in every relevant aspect; you, in fact, suspect that they are not. For example, you suspect that, unbeknown to Urnie, balls in one color group (you don't know which group) are too small, so that they go to the bottom of the urn in a way that they have a very low chance of being picked (e.g. if white balls are the small ones, despite the fact that they constitute 90% of the balls inside the urn, picking them would become very improbable, say, with the chance of only one in 100, when the chance of a black ball being picked is 99 in 100) (hence, the second criterion is not satisfied).

In the *Fair Urn* case, we can be confident that the probability of choosing a white ball is the proportion of white balls in all the balls (i.e. 90%) and we can confidently make the judgment that the picked ball is a white ball based on such statistical evidence. In contrast, in the *Unknown Urn* case, we are not confident about the randomness of the selection process or the fact that balls have the same chance of being chosen. This casts serious doubts on the relevance and applicability of the statistical evidence—which relies on the proportion of white balls to all balls—and decreases our overall confidence in (and increases the margin of error for) any judgment based on such statistical evidence.

It is worth noting that if we *have to* choose between two colors in the *Unknown Urn* case (say, someone will shoot us if we don't!) or we really want to bet on one of the colors, it would be more rational to go for the white color rather than the black, given the only available evidence that the white balls are in the majority. However, such choice fails to meet the high standards required for grounding knowledge or conviction since there is very low confidence regarding its truth and it has a huge margin of error. In fact, if we are not forced to make a choice, we should seriously consider suspending the judgment in cases like *Unknown Urn*. In short, the more we are confident that the aforementioned criteria obtain, the more we would be confident and justified in making a judgment on the basis of the available statistical evidence; and the less we are confident about those criteria, the less we should be willing to make any judgment. That said, in the actual situations, any choice should be made in the light of the overall margin of error as well as many other considerations like what is at stake if we suspend judgment, make a right/wrong choice, etc.

IV.3.3 Discussing the Use of Statistical Evidence in Some Cases from the Literature

In what follows, I discuss some of the famous cases of statistical evidence from the literature and examine the margin of error for each one with a particular emphasis on the aforementioned criteria for minimizing the second type of error.

Let’s start with the previous *Smith v. Red Cab* case. Thomson concludes that “[i]f we believe Mrs. Smith's story, and are aware of no further facts that bear on the case, then we shall think it .6 probable that her accident was caused by a cab operated by Red Cab” (Thomson, 1986. p.199). She explains that relative to the facts reported by (1) “The cab which caused the accident was a cab in town that night”; and (2) “six out of ten of the cabs in town that night were operated by Red Cab” we conclude that: (3) “The probability that the cab which caused the accident was operated by Red Cab is .6” (Thomson, 1986. p.199).

Now, as far as the third type of error (namely, the probability of error given the evidence is plausibly applicable to the case) is concerned, the evidence seems to be acceptable in this case (it seems more probable than not, that the cab which caused the accident was operated by Red Cab). But what about two other types of error?

The first type of error pertains to the error in the background data of the case like (1) and (2). Arguably, part of the uncomfortableness we feel at the idea of basing conviction on such statistical evidence about Red Cab’s market share is due to the fact that, in this case, even a small amount of the first type of error can be devastating for the putative conclusion. For example, provided there is only 20% chance that the car which caused the accident is not indeed from any of these two cab companies (either it is not a cab at all—which is not much unlikely given it was dark, etc.—or it is a cab from another town), the overall chance that Red Cab company is behind the accident drops to 0.48[10] which makes it more improbable than not that the Red Cab is, in fact, culpable (hence fails to satisfy the required *preponderance of evidence* standard*)*.

Thomson (1986) herself recognizes that the weight we place on the hypothesis that it was a cab that caused the accident is affected by, say, the number of non-cabs around that are disguised as a cab. However, she doesn't discuss the effect of such possibility of error on the putative overall probability that she assigns to Red Cab being behind the accident and just supposes that we've believed Mrs. Smith's story and assumed that the background information is correct (that is, the first type of error is zero or infinitesimal). In any case, the following *Blue-Bus* case—which is essentially very similar to the *Smith v. Red Cab* case—by decreasing the probability of the third type of error to 10% (instead of 40% in the *Smith v. Red Cab* case) can decrease these worries:

> *Blue-Bus*
>
> A bus causes harm on a city street. Suppose there are no witnesses to the incident, but we have evidence that 90% of the buses operating in the area, on the day in question, were owned by the Blue-Bus company (Smith, 2018, p.1195).

Putting aside the worries about the first type of error, I will now focus on the second type of error. To make sure that the second type of error with respect to such probabilistic judgments is low, firstly, we need to have good reasons to think that the fact that a bus from Blue-Bus company caused the harm on the city street (or a cab from the Red Cab company caused the accident) has obtained randomly. For that, we need to be confident that none of

[10] Call P(Red Cab company being responsible) P(r) and P(The car behind the accident is from one of two cab companies) P(t): Assuming P(r|t) = 0.6 and P(t) = 0.8 (given the aforementioned 20% chance of error), given the *product rule*: P(r & t) = P(r|t) P(t) therefore: P(r & t) = 0.6 * 0.8 = 0.48; now since P(r) = P(r & t)/P(t|r) and given P(t|r) = 1, we conclude: P(r) = 0.48/1 = 0.48

the companies or drivers, predicting the results of such probabilistic judgment, intentionally caused the accident, say, to smear the other company.

Moreover—regarding the second criterion—one may argue that we cannot be confident that the probability of a bus from Blue-Bus company being the cause of the accident is the proportion of the buses from Blue-Bus in all buses (i.e. 90%) since this requires us to be confident that buses from both companies have a similar chance of being the cause of the accident while there is a good chance that two companies differ on various factors like their driver's average age, average years of experience, previous records, etc., that can significantly affect the chance of having an accident (e.g. what if Blue-Bus has no record of accident for the last 10 years, while the other company has several accidents every year).

In drawing the conclusion regarding the *Smith v. Red Cab* case, Thomson explains:

> [T]hose [(1) and (2)] are the only facts such that we are both aware of them and aware of their bearing on the question who operated the cab which caused the accident. (Perhaps we are aware that the accident took place on, as it might be, a Tuesday. Even so, we are not aware of any reason to think that fact bears on the question whose cab caused the accident.) Other facts whose relevance is clear might come out later: for example, a Green Cab driver might later confess. But as things stand, we have no more reason (indeed we have less reason) to think that any facts which later come out would support the hypothesis that the cab which caused the accident was operated by Green Cab than we have to think they would support the hypothesis that the cab which caused the accident was operated by Red Cab. We are therefore entitled to conclude that (3) is true--in fact, rationality requires us to conclude that (3) is true, for .6 is the degree of belief that, situated as we are, we ought to have in the hypothesis that the cab which caused the accident was operated by Red Cab (Thomson, 1986, p. 200).

Here Thomson seems to admit that there might be some relevant factors in support of either of the hypotheses that may come out later, but she rightly explains that as things stand, we have no reason to think such future evidence would support say the first hypothesis rather than the second. In other words, as far as we know, there is no reason to prefer one hypothesis over the other in virtue of anything other than the available evidence (i.e. the market share). However, I don't agree with Thomson that we are "entitled to conclude" that the probability that the Red Cab company is responsible for the accident is 0.6 or we "ought to have" 0.6 degree of belief in this hypothesis. Although as she rightly says, such considerations about unknown facts cannot be used in favor of any of the two hypotheses, they can be used in favor of a third option, namely, suspending the judgment[11].

This depends on many details from the actual situation, but in case we are not confident that the process is random or cars from each company have similar chances of being the cause of the accident (i.e. the margin of error is high), why do we need to make any judgment in the first place? In other words, when there is a good chance that the statistical evidence is irrelevant or the putative probability is inaccurate, why should we proceed to draw any conclusion on that basis, instead of suspending the judgment due to lack of sufficient evidence? (again, the assumption is that no one will shoot us if we don't opt for one of the two hypotheses!).

Accordingly, the *Smith v. Red Cab* and *Blue-Bus* cases are more like *Unknown Urn* case than *Fair Urn*. The case below illustrates a modified version of *Blue-Bus* case which makes it closer to acceptable cases like *Fair Urn*:

> *Autonomous Buses*
>
> All the cars, including all the buses, in the city, are fully autonomous with no driver in charge. There are two companies in the city that own all the buses: Silver-Bus and Gold-Bus. All the buses have been manufactured in the same factory and are identical in every aspect except their color. A bus causes harm on a city street. Although there are no eyewitnesses to the incident, we have very good evidence that the bus belongs to one of the two bus companies in the city and that of all the buses operating in the area in that period, 90% were owned by the Gold-Bus company and other 10% belonged to the Silver-Bus company. We also have good evidence that no one hacked into the system.

In the *Autonomous Buses* case, we seem to have more confidence that the accident was a random occasion, and the buses from each company have a similar chance of being the cause

[11] Just like the case of belief, court verdicts can be regarded as a threefold matter: An affirmative verdict is like believing p (i.e. the defendant is guilty) and a negative verdict covers both believing not p (i.e. the defendant is not guilty) or suspension of judgment; this threefold structure is more clear, for example, in criminal cases in Scots law in which in addition to the verdicts of "not guilty" and "guilty" there is a third verdict of "not proven" (Smith, 2018, p. 1203).

of the accident. Accordingly, in this case, we are more confident to think that the probability of a bus from Gold-Bus company being responsible for the accident is quite high and thus imposing liability on Gold-Bus company on the basis of such statistical evidence has a lower margin of error compared with the Blue-Bus case[12] (and I think many people are more comfortable with it).

My main point here is that the verdict about many statistical cases requires caution and is not as straightforward as it is sometimes presented in the literature. The decision about whether we can base judgment on statistical evidence in each case depends on many factors like our confidence about the satisfaction of some criteria that are not themselves statistical matters and many actual details of the cases including details about what is at stake in each side. In criminal cases, what is at stake of making a wrong judgment is very high and the relevant standard of proof—i.e. *beyond reasonable doubt*—is much more demanding (usually quantified as more than 90-95% confidence). For example, take the following case:

> *People v. Tice*
>
> Two people, Tice and Simonson, both hated Summers and wished him dead. Summers went hunting one day. Tice followed with a shotgun loaded with ninety-five pellets. Quite independently, Simonson also followed, but he had loaded his shotgun with only five pellets, that being all he had on hand. Both caught sight of Summers at the same time, and both shot all their pellets at him. Independently: I stress that there was no plot or plan. Only one pellet hit Summers, but that one was enough: it hit Summers in the head and caused his death. While it was possible to tell that the pellet which caused Summers' death came either from Tice's gun or from Simonson's gun, it was not possible to tell which (Thomson, 1986, pp. 200-201).

Let's assume that the first type of error for this case is infinitesimal. Apparently, the third type of error is very low too (assuming the evidence is plausibly applicable to the case, there is only 5% error in making the judgment that Tice is the killer). Thomson claims that in this case we "should" conclude that it is true that "[t]he probability that the pellet that caused

[12] Maybe a better approach to this case—that prevents from the conviction of the company with the higher market share in every similar future incident—is to ask the companies to share the costs and pay for the harm in accordance with their market share.

Summers' death was fired by Tice is .95" (therefore the fact that Tice killed Summers is beyond reasonable doubt) (Thomson, 1986, p.201). However, this is only correct if the second type of error is infinitesimal.

The second type of error would be sufficiently low just in case, first, we are confident that the fact that a pellet from Tice rather than Simonson caused Summers' death is obtained randomly. This requires us to have good reasons to think that both Tice and Simonson were similarly willing to kill Summers and put the same effort in it, and therefore the fact that one of them managed to do that was a matter of chance. But what if while, say, Simonson was really serious in targeting Summers, Tice didn't intend to kill Summers (or changed his mind in the very last moment) and somehow intentionally avoided from targeting Summers (though shot all his 95 pellets at him to scare him)? Thomson seems to just assume that there is no such possibility when she says: "If Tice did not cause Summers' death, then his failure to do so was--relative to the evidence we have in hand--just luck… He did everything he could to cause the death" (Thomson, 1986, p.201). However, I'm not sure the way the case is described allows us to confidently rule out the possibility that Tice didn't put as much effort as Simonson or didn't really have the intention to kill Summers—which, in turn, can make the statistical evidence irrelevant and probably useless.

Secondly, we need to be confident that a pellet from Tice had a similar chance of killing Summers as a pellet from Simonson. However, there is a good chance that pellets from each of them had, in fact, different chances of hitting Summers due to various factors like the difference in their shooting skills, etc. A skilled shooter may be able to hit the target with a few pellets, while a bad shooter may shoot a hundred pellets none of which hits the target. Accordingly, I think factors like the skills of the shooters (which are absent from the available body of evidence) have much more relevance and significance in determining who killed Summers and merely focusing on the number of pellets is misleading.

Note that in *People v. Tice* case, even a very low margin of error including a small margin of the second type of error (namely, a small chance that the fact is not obtained randomly or pellets from each of them have different chances of hitting Summers) may drop the degree of confidence in the proposition that Tice is the killer from the current 95% in a way that it wouldn't meet the *beyond reasonable doubt* standard of proof for criminal cases anymore. That being so, if there is not much confidence that the required criteria are satisfied, there shouldn't be much confidence in basing a conviction on such statistical evidence as well.

The following is a modified version of *People v. Tice* in which the probabilistic judgment is based on similar statistical evidence, yet there is high confidence about the judgment since the required criteria are met: Imagine instead of a shotgun they use a shooting machine that is fixed somewhere on the ground. All they need to do is to lock on the target and press the fire button once. The available evidence suggests that they both have locked on the target and pressed the button and the guns have started shooting at the same time. The two guns are from the same model and shoot the target with a similar standard deviation. The only relevant difference between these guns is that Tice's gun had 95 while Simonson's gun had only 5 pellets.

That said, the application of statistical evidence is not confined to such unrealistic scenarios. For example, consider the *Joe and television* case that I mentioned in this thesis' introduction. Assume we are very confident about the background information in this case (i.e. the number of stolen TVs, etc.), that is, the first type of error is infinitesimal. The third type of error, in this case, is also apparently very low (i.e. 1%). So, I'll get to the second type of error. First, we need to be confident that the police apprehended one of those 100 persons randomly and did not have the intention of capturing a *specific* person (in this case Joe). Secondly, we need to be confident that an innocent and a guilty person both had a similar chance of being captured (e.g. it was not the case that the innocent person was slower and thus had a higher chance of being captured).

Such confidence depends on the details about the case and the process of apprehension, but I think we can be quite confident about the satisfaction of these criteria in many normal situations. Just in case, imagine that the policeman was in a position to apprehend any of those 100 people, yet he only had enough time to capture one of them; so, he came up with a random number between 1 and 100 and went for the person with that number. In this case, we can be pretty confident that the required criteria are met.

That said, the observant reader may ask what is the difference between factors like average years of experience in *Blue-Bus* case or the shooting skill in *People v. Tice*, on the one hand, and Joe's character, criminal record, financial situation, etc., in *Joe and television* case, on the other? Why don't such factors similarly increase the margin of error in *Joe and television* case?

In response, there is, in fact, a crucial difference between *Joe and television* case and previous cases. Recall that what led to such increase in the second type of error in previous cases was that those factors could affect the chances of, say, a bus from other company rather than Blue-Bus cause the accident, or a pellet from Simonson rather than Tice kill Summers. In other words, in previous cases, there was a good chance that due to such factors members of the population do not have the same chances of being chosen which is a requirement for taking the probability to be the proportion of those items in the population (e.g. number of buses from Blue-Bus in all buses).

In contrast, in the *Joe and television* case factors like Joe's criminal record or financial situation or even him being indeed innocent or guilty do not seem to affect the chance of him—among 100 people with TV—being apprehended. Accordingly, unlike the previous cases, in *Joe and television* case, we are more confident that each member of the (100 people) population—either an innocent person or a guilty one—has an equal chance of being chosen (apprehended) and thus the probability of the apprehended person being indeed a thief is the proportion of the TV thieves to all people who walked out of the store carrying TVs (i.e. 99%). This allows us to draw a parallel with an acceptable case of relying on statistical evidence like the previous *Fair Urn* case or an ordinary case of random selection like the following:

a) There are 100 people carrying TVs, 99 of them are thieves and 1 is innocent (there are 100 marbles in the bowl, 99 black and 1 white)
b) One *random* person from the 100 people is apprehended (1 marble is randomly picked)

Therefore:

c) There is a 99% chance that the apprehended person is a thief (there is a 99% chance that the marble is black)

Note that further information about Joe's criminal records, etc. can, of course, help to decrease the chance of error even further, but crucially even without any extra information and only based on the facts reported by (a) and (b) we can confidently conclude (c) and be quite sure that the chance of making a wrong judgment, in this case, is sufficiently low.

At last, it is worth noting that the provided criteria for applicability of statistical methods can help to address some difficult cases in this context: On the one hand, it sheds light on the problem with making judgments based on statistical data about race, gender, ethnicity, etc.; on the other hand, it gives a clue as to why despite general reluctance of the courts to base conviction on statistical evidence, cases like DNA evidence which are essentially a kind of statistical evidence are usually accepted by courts (Smith, 2018, p. 1213).

Regarding the former, imagine in the *Smith v. Red Cab* case, the police have managed to narrow down the list of suspects to two specific drivers with heterogeneous characteristics (say, they are from two different genders, races, etc.). Suppose, we don't have any evidence to decide which one is indeed behind the accident other than the statistical data that shows the people with the same, say, gender as the first driver (call it gender A) have twice chance of causing an accident compared with people with the same gender as the second driver (call it gender B). Now, given this is all the evidence we have, it may be tempting to conclude that *indeed* the probability of the first driver being behind the accident is twice as the second driver (i.e. 66% compared with 33%). However, as the previous discussions show this conclusion is implausible and misleading. This is because, again, even if we assume that such statistical data has some relevance, without having good reasons that the required criteria are satisfied, namely without having the confidence that the accident has happened randomly and in absence of information about many other relevant factors that can affect the chances involved (e.g. perhaps the first driver, unlike the second driver, is very experienced, or maybe, despite the fact that people from gender A on average cause more accidents, only those of them opt for taxi driving that are exceptionally good at driving, etc.), any conclusion will have a high margin of the second type of error and would be indefensible.

Regarding the case about DNA evidence, take the following example:

> Suppose a DNA sample is lifted from a crime scene and identified as belonging to the perpetrator. When run against a database the sample yields a 'cold hit' – a matching DNA profile belonging to some member of the population, with no other known connection to the crime. Suppose this individual is arrested and charged with the crime. Suppose that no further evidence against the individual emerges, but an expert witness testifies that the chance of two individuals in the population sharing the same profile is exceedingly slim[13] (Smith, 2018, p. 1212).

Smith refers to this case as a difficulty for his account because while courts rely on such evidence it does not satisfy the criteria suggested by Smith about providing normic support

[13] A conviction based on a cold hit DNA may be wrongful not only if there is some laboratory error, etc., (sometimes called a "false" match) but crucially even if there is no error and match is "true" (Smith, 2018, p. 1212).

(because the situation in which, despite the DNA match, the unlucky arrested person is, in fact, innocent—since the sample at the crime scene belongs to the other person who matches the profile—does not seem to require special explanation) (Smith, 2018).

However, such evidence can be seen as a case which apparently satisfies the provided criteria in this chapter. Assuming the background information is correct (i.e. the first type of error is infinitesimal), we can be confident that the second type of error is sufficiently low. This is because in normal situations, first, there is high confidence that the DNA match was obtained randomly (i.e. no one in the laboratory has deliberately changed the results) and, second, there are good reasons to think that other factors like the defendant's criminal records, motives, character, etc., and even her being indeed guilty or innocent do not affect her chance of getting a wrongful cold hit DNA result (i.e. items in the population have the same chance of being chosen—that is, being the one that gets a wrong match). Finally, the third type of error is also infinitesimal[14]. That being so, we have good reasons to think that the overall margin of error, in this case, is sufficiently low and, thus, we can confidently base a conviction on such evidence.

IV.4 CONCLUSION

If applied correctly, statistical methods can produce pieces of evidence that lead to extremely reliable judgments. We have seen that in science; why not elsewhere? Even so, the statistical judgments are especially vulnerable to various kinds of error, and we are prone to commit many known and unknown biases with respect to statistical data; thus, applying statistical methods especially in critical situations like legal cases demands extreme caution.

[14] The chance that two persons share a DNA match can be as low as 1/400 trillion (Gardiner, forthcoming, p. 6).

In this chapter, I argued that many famous cases in the literature that are supposed to provide an example of a judgment based on statistical evidence to be juxtaposed with non-statistical cases may, in fact, be instances of problematic application of statistical methods. Accordingly, although the putative conclusion in cases like *Smith v. Red Cab*, *Blue-Bus* and *People v. Tice* seems problematic, it is not because, in general, statistical evidence cannot be an acceptable basis for judgments; instead, the problem can be traced back to the lack of confidence about the satisfaction of the required criteria for application of statistical methods in those cases which in turn leads to a high margin of error pertaining to the conclusion. That being so, in many other cases like *Autonomous Buses*, *Joe and television* and DNA evidence in which we are confident that the required criteria are met, we can plausibly and confidently base judgments on the statistical evidence.

SUMMARY

Statistical methods can produce valuable predictions and extremely reliable judgments either in epistemology or the courtroom even in situations where other techniques fail. However, we feel uncomfortable at the idea of basing knowledge or conviction on statistical evidence. Given the great reliability of statistical evidence, on the one hand, and the failure of attempts to provide a rationale for discriminating against it, on the other, and in the light of historical considerations that can help to explain our harsh intuitive reactions to it, there is a good case for the claim that despite its counter-intuitiveness statistical evidence can base knowledge and conviction.

The 2nd and 4th scenarios in chapter 1 illustrate some cases of belief based on statistical evidence which can be considered as cases of knowledge in the framework of popular accounts of knowledge like safety. However, the knowledge assignment to cases like Lottie requires a decrease in commitments to the intuitive notion of knowledge which is, arguably, the product of various historical factors. This allows moving towards more scientific accounts of knowledge which can also help with addressing some long-lasting issues in epistemology. Furthermore, a similar departure from intuitions allows us to defend the use of statistical evidence in legal cases, at least as far as the justificatory status of such evidence is concerned.

That said, there are serious concerns about the way the statistical methods are applied in real situations. Arguably, even some famous cases from the literature on statistical evidence may be dubious applications of statistical methods. Thus, although statistical evidence can be a basis for knowledge and conviction, it requires great caution (and even training) about potential biases and fallacies, and it demands high confidence that the required criteria are met and the overall margin of error is sufficiently low. Maybe such confidence together with more familiarity with statistics, bring more comfortableness pertaining to statistical evidence to the table.

REFERENCES

Bargh, J. and Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), pp.462-479.

Crane, T. (1995). ‘Causation’, ‘Time’ and ‘Universals’. In: A. Grayling, ed., *Philosophy: A Guide Through the Subject*. Oxford: Oxford University Press, pp.184-214.

Csibra, G. and Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), pp.148-153.

Enoch, D., Fisher, T. and Spectre, L. (2012). Statistical evidence, sensitivity and the legal value of knowledge. *Philosophy and Public Affairs*, 40(3), pp.197-224.

Gardiner, D. [Forthcoming]. Legal Burdens of Proof and Statistical Evidence. *Routledge Handbook of Applied Epistemology.*

Goldman, A. (1967). A Causal Theory of Knowing. *The Journal of Philosophy*, 64(12), pp.357–372.

Hacking, I. (2006). *The Emergence of Probability*. 2nd ed. New York: Cambridge University Press.

Innocence Project. (2017). *DNA Exonerations in the United States - Innocence Project*. [online] Available at: fhttps://www.innocenceproject.org/dna-exonerations-in-the-united-states/ [Accessed 17 Dec. 2018].

Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Mackie, J. (1965). Causes and Conditions. *American Philosophical Quarterly*, 2(4), pp.245–264.

Nozick, R. (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

Orban, G., Fiser, J., Aslin, R. and Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, 105(7), pp.2745-2750.

Papineau. D. [Forthcoming]. The Disvalue of Knowledge.

Piantadosi, S., Tily, H. and Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), pp.3526-3529.

Pritchard, D. (2007). Anti-Luck Epistemology. *Synthese*, 158(3), pp.277-297.

Pritchard, D. (2012). Anti-Luck Virtue Epistemology. *Journal of Philosophy*, 109(3), pp.247-279.

Pritchard, D. and Smith, M. (2004). The psychology and philosophy of luck. *New Ideas in Psychology*, 22(1), pp.1-28.

Sellars, W. (1963). Philosophy and the Scientific Image of Man. In: K. Kevin and R. Robert B., ed., *In the Space of Reasons: Selected Essays of Wilfrid Sellars*. Cambridge: Harvard University Press.

Simon, R. (1969). Judges translations of burdens of proof into statements of probability. *Trial Lawyer’s Guide* ,13, pp.103-114.

Smith, M. (2018). When Does Evidence Suffice for Conviction? *Mind*, 127(508), pp.1193-1218.

Sosa, E. (1999). How to Defeat Opposition to Moore. *Nous*, 33(13), pp.141-153.

Tenenbaum, J., Kemp, C., Griffiths, T. and Goodman, N. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, 331(6022), pp.1279-1285.

Thomson, J. (1986). Liability and Individualized Evidence. *Law and Contemporary Problems*, 49(3), pp.199–219.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), pp.293-315.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, T. (2009). In: P. Greenough and D. Pritchard, ed., *Williamson on Knowledge*. Oxford: Oxford University Press, pp.279–384.

Zipf, G. (1949). *Human behavior and the principle of least effort*. New York: Addison-Wesley.