Dimensional Reliabilism

S. ORESTIS PALERMOS
University of Ioannina

The paper argues that (i) the notion of epistemic reliability, as it is standardly defined within mainstream epistemology, is a multidimensional concept, and that (ii) paying attention to reliability's multidimensional nature can significantly expand reliabilism's purview both in the theoretical and practical domain. Reliabilist theories of knowledge and justification agree that a process is reliable just in case it leads to a 'sufficiently high preponderance of true over false beliefs.' Given this straightforward definition, reliability appears to be, and so far has been implicitly treated by epistemologists as, a mono-dimensional concept. Nevertheless, the terminology and formulas that statisticians and computers scientists have developed for assessing the reliability of binary classification processes reveal that for any belief-forming process whose aim is to identify the presence or the absence of some quality ι, there are at least four different interpretations of the above definition. These interpretations correspond to four separate dimensions of reliability: Precision, Negative Predictive Value (NPV), Recall and Specificity. Though epistemologists do not explicitly specify how they interpret the above definition of reliability, due to their main interest in 'know-that', their mono-dimensional focus seems to have so far been mainly directed at Precision alone (or, at best, to both Precision and NPV—but without distinguishing between the two). As it transpires, however, the (entirely) neglected dimensions of Recall and Specificity can guide the satisfaction of another important epistemic goal, best captured by the locution 'know-most/all.' The upshot is that different epistemic goals call for high levels of reliability along different dimensions (and, often, their combinations too). This observation invites upgrading reliabilism into *dimensional* reliabilism, according to which assessing whether a process is knowledge-conducive depends on whether it manifests high reliability along the appropriate dimension(s), given the epistemic goal at hand.

1. Introduction

In everyday contexts, epistemic reliability is widely used to assess the trustworthiness of several information sources: Informants, medical tests (e.g., SARS-COV-2

Contact: S. Orestis Palermos <sopalermos@uoi.gr>

PCR and rapid antigen testing),[1] newspapers, blogs, websites, directory services, TV programmes, social media, intelligence agencies, sense modalities—the list is as long as the information channels people rely on. Over the past few decades, reliability has also become a central concept within the theoretical discipline of mainstream epistemology. Generic reliabilism and related views such as process reliabilism and virtue reliabilism assume that knowledge is true belief that is the product of a *reliable* belief-forming process or cognitive ability. While considerable disagreement exists between these views on what more is required from a process to count as knowledge-conducive, they all put forward the same definition of reliability:

> *Reliability*
> A process is reliable just in case "it leads to a sufficiently high preponderance of true beliefs over false beliefs" (TB >> FB).[2]

Given this straightforward definition, reliability appears to be, and so far has been implicitly treated by epistemologists as, a mono-dimensional concept. Nevertheless, by reference to widely used functions that statisticians have developed for assessing the reliability of binary classification processes, this paper argues that reliability is a *multidimensional* concept. Specifically, given the above definition of reliability, the notion is open to at least four different interpretations—what (following statisticians) we will here refer to as 'Precision,' 'Negative Predictive Value' (NPV), 'Recall' and 'Specificity.' Moreover (as the argument will go) each of these dimensions of reliability serves different epistemic goals and, depending on the agent's epistemic goal, they can all take centre stage in assessing whether a belief-forming process is reliable, and thereby knowledge-conducive.

Pondering in this way on the multidimensional nature of reliability will bring to the fore two broad epistemic goals that correspond to two important kinds of knowledge (or epistemic states, if you like). The first goal is best captured by the familiar locution of 'know-that.' Though epistemologists do not specify how they interpret the above definition of reliability, due to their interest in 'know-that', their mono-dimensional focus seems to have been mainly directed at Precision alone (or, at best, at both Precision and NPV, but without distinguishing between the two).[3]

---

1. 'PCR' is the widely used acronym for 'quantitative reverse-transcriptase polymerase chain reaction' testing. SARS-CoV-2 is the virus responsible for COVID-19.

2. The quote comes from Sosa (1992, p. 79). See also Goldman (1979), Goldberg (2010), Goldman and Beddor (2021).

3. In the literature, the most indicative example in support of the claim that epistemologists may have mainly focused on Precision alone comes from the way Peter Graham (2014a) responds to Tyler Burge's 'rabbits counterexample' against the claim that correct representation is a biological function. To respond to Burge, Graham goes into sufficient detail about his thoughts on reliability in a way that seems to clearly suggest that he prioritises Precision. I extensively comment on this insightful discussion, in Section 7.

As it will transpire, however, the (entirely) neglected dimensions of Recall and Specificity are essential for guiding the satisfaction of another kind of epistemic goal, best captured by the locutions 'know-most' and 'know-all.'[4] Think, for example, of sentences such as 'I know all movies by David Lynch'; 'I know most subjects who have contracted the virus'; 'I know most/all profitable investments'; and so on.

Though off the radar of reliabilist epistemology so far, the paper argues that 'know-most/all' and the dimensions of reliability that distinctively underly it—that is, high Recall and/or high Specificity—are important epistemic concepts that epistemologists should not neglect anymore. Various reasons can be offered in support of this claim: For example, manifesting high Recall and/or high Specificity is essential for alleviating a common kind of ignorance (to be specified later); from a reliabilist perspective, the associated epistemic states of 'know-most/all' seem irreducible to 'know-that'; focusing on Recall and Specificity can provide significant insights concerning the debate over the validity of evolutionary reliabilist epistemology; and, perhaps most importantly, paying attention to Recall and Specificity can lead to useful theoretical and practical distinctions for assessing and understanding within epistemology the reliability of various information sources (including informants, medical tests and mass media)—distinctions, which are impossible to draw on the basis of a mono-dimensional approach to the concept of reliability.

For example—to appeal to an until recently topical distinction—we will see that if reliabilists only focus on the dimension of Precision and/or Negative Predictive Value, without any recourse to the dimensions of Recall and Specificity, they are unable to explain the most noteworthy and crucial difference between the reliability of PCR testing and the reliability of (at least certain) rapid antigen tests for SARS-CoV-2. And yet, as it is widely known, there is often supposed to be an important difference between the reliability of these two kinds of medical tests for SARS-CoV-2. As we shall see, the present approach, with its explicit recognition of the multidimensional character of the notion of reliability, can provide reliabilism with the means for clearly accounting for this distinction—and many more besides.

Overall, then, the aim of the following is to highlight how each of the aforementioned dimensions of reliability can serve different epistemic goals and that, depending on the agent's epistemic goal, they can all take centre stage in assessing whether a belief-forming process is reliable, and thereby knowledge-conducive. This observation, I will argue, invites us to upgrade reliabilism into what I will refer to as *dimensional* reliabilism.

To make the case for the above, the paper will proceed in the following order. To first familiarise us with the relevant dimensions of reliability, the paper will

---

4. I take it that that the two expressions refer to epistemic states that satisfy the same *kind* of epistemic goal, but to different *degrees*. That is, the epistemic state captured by the expression I 'know all *p*s' is an idealised version of the epistemic state captured by the expression I 'know most *p*s'.

start with a nontechnical discussion of the target dimensions (Section 2), which will then be followed by a section offering mathematically rigorous definitions of them, drawing on widely used statistical tools (Section 3). With these definitions in place, the paper will then proceed to provide several examples that illustrate the central idea—that is, that different epistemic goals call for high scores across different dimensions of reliability and their combinations (Section 4). At that point, the discussion will then be paused to consider a technical worry: The dimensions of Precision and NPV seem to directly relate to 'know-that', whereas Recall and Specificity relate to 'know-most/all'; could there be a way to reduce Recall and Specificity to Precision and NPV, such that there would be no need for epistemologists to look beyond 'know-that'? Arguing that this seems impossible (Section 5) and that 'know-most/all' is not thereby reducible to 'know-that' (at least not from a reliabilist perspective), the paper will continue to highlight some of dimensional reliabilism's key advantages: The view is in an excellent position to explain, in reliabilist terms, how knowledge can be the antidote to a very common kind of ignorance (Section 6). It can also prove particularly helpful in defending reliabilist approaches to evolutionary epistemology from a worrying counterexample (viz., the 'rabbits counterexample') (Section 7). Finally, the paper (Section 8) will demonstrate how the view offers the tools for a more nuanced understanding of the reliability of information sources, both in the theoretical domain (for example in our assessment of thought experiments) and in practice: For instance, dimensional reliabilism is in a unique position to explain the difference in the reliability of medical tests such as PCRs and at least certain rapid antigen tests for SARS-CoV-2 as well as to offer a detailed (and so far inaccessible) breakdown of the several ways in which online information sources and mass media might succeed (or fail) in being reliable.

2. Dimensions of Reliability

Let's start by familiarising ourselves with the several dimensions of reliability with a simple example. Think of a process whose function, say, is to identify cats—that is, a cat-identifying mechanism (C). For many, claiming that C is reliable will likely indicate the following:

1. If C outputs that an object is a cat, then it is highly likely that the target object is a cat.

This kind of reliability can be grasped in terms of a 'high preponderance of true over false beliefs'—that is, in terms of the aforementioned, widely accepted definition of reliability—in the following way:

> True beliefs of the type 'object x is a cat' >> False beliefs of the type 'object x is a cat'

Indeed, this is an important way in which C can be reliable. It suggests that the mechanism will rarely misidentify objects that are not cats as cats. Nevertheless, there are further cat-related beliefs C is responsible for that are not guaranteed to be equally reliable:

2. If C outputs that an object is not a cat, then it is highly likely that it is not a cat.

This kind of reliability can be grasped in terms of a 'high preponderance of true over false beliefs' in the following way:

> True beliefs of the type 'object x is not a cat' >> False beliefs of the type 'object x is not a cat'

A mechanism that satisfies the first kind of reliability may not satisfy the second. For example, even if C is highly unlikely to misidentify objects that are not cats as cats, there might be lots of (strange-looking) cats that the mechanism mistakes for objects that are not cats.

3. C will detect most existing cats. For objects that are cats, C's outputs that the target objects are cats are considerably higher than C's outputs that the target objects are not cats.

Interestingly, this kind of reliability can also be grasped in terms of a 'high preponderance of true over false beliefs' in the following way:

> True beliefs of the type 'object x is a cat' >> False beliefs of the type 'object x is not a cat'

Again, it is easy to see how a mechanism that satisfies the first kind of reliability may fail to satisfy the third. There might be several cats in the vicinity, but the mechanism fails to categorise them as cats (because, say, they are strange-looking cats). As a result, a large percentage of cats will be ignored.

4. C will detect most objects, which are not cats. For objects that are not cats, C's outputs that the target objects are not cats are considerably higher than C's outputs that the target objects are cats.

This too can be grasped as a kind of reliability in terms of a 'high preponderance of true over false beliefs' in the following way:

> True beliefs of the type 'object x is not a cat' >> False beliefs of the type 'object x is a cat'

Though this case might be slightly trickier, it is not difficult to see how a mechanism that satisfies the first kind of reliability may fail the fourth kind. It could be the case that, in the vicinity, the number of cats is much larger than the number of objects that are not cats. In this case, the number of objects that are not cats (and thus the potential number of correctly identified such objects) will be close to the number of objects that are incorrectly identified as cats, even if—due to the first kind of reliability—this happens only rarely. In this scenario, true beliefs of the type 'object x is not a cat' will not be much higher than false beliefs of the type 'object x is a cat'. In result, of all the objects that are not cats, a large percentage will be ignored.

The upshot is that defining reliability of a mechanism as the mechanism's ability to lead to a 'preponderance of true over false beliefs' is too broad. According to the above, for any kind of epistemic belief-forming process whose function is to identify the presence of some quality ι (such as the quality of being a cat), there are at least four distinct ways to interpret mainstream epistemology's generic definition of reliability, thereby revealing four different dimensions of reliability. What then is the way to decide which of the above interpretations is the correct one or whether any single one of them is correct? Put another way, which 'sufficiently high preponderance of true over false beliefs' should assessments of whether a belief-forming process is reliable, and thus knowledge-conducive, focus on? Could there be contexts where one should be preferred over the other?

3. Reliability Indicators

More precise definitions of the above dimensions of reliability can be provided by employing certain widely used functions that statisticians and computer scientists have developed for assessing the reliability of classification processes (see, for example, Tharwat 2021). Binary classification can be defined as a kind of categorisation process, whereby sampled items are divided into two distinct classes, depending on whether the target items bear or lack some qualitative property ι. As we go along we will explore various examples that demonstrate the potential practical applications of binary classification, but the following are common instances where this kind of categorisation process can be implemented in practice:

- Image classification: Assessing whether an image represents a certain visual pattern (e.g., a cat) (where ι is the quality of representing the target pattern).
- Spam detection: Assessing whether a message is a spam (where ι is the quality of being a spam).
- Medical screening: Assessing whether a subject has a target condition or pathogen (where ι is the quality of having the target condition or pathogen).

Binary classification is relevant to epistemology, because most (perhaps all) basic human belief-forming processes seem amenable to be construed as binary classification processes. Consider the following examples:

- Visual processes of colour perception (where ι is a specific colour).
- Visual processes of pattern recognition (where ι is a specific pattern).[5]
- Long term autobiographical memory (where ι is a past event in one's life as opposed to confabulation).
- Long term explicit memory (where ι is a previously encoded piece of information as opposed to confabulation).
- Inference-making (where ι is a valid inference).
- Reception of testimony (where ι is a reliable testimony).[6]

Accordingly, the reliability of most human belief-forming processes can be assessed on the basis of classification functions.[7] To understand how classification functions work as indicators of different kinds of reliability, the following terminology will be useful (see also (Tharwat 2021) for more details):

For any set of elements that might bear the property ι (e.g., being a cat):

- Positives, P, refer to the tested elements that bear ι (e.g., to actual cats).
- Negatives, N, refer to the tested elements that do not bear ι (e.g., to any object that is not a cat).
- True Positives, TP, refer to the elements that the test identified as bearing ι and they bear ι indeed (e.g., to those elements that they were identified as cats and they are cats indeed).

---

5. Hearing, smell, touch, and taste can all be modelled in a similar way.

6. It is here assumed that testimony is a basic source of knowledge. If it is not, then it could be explained by some appropriate combination of basic belief-forming processes.

7. The generality problem (Feldman 1985; Conee & Feldman 1998) is orthogonal to employing this approach for assessing the reliability of belief-forming processes. However one chooses to type the relevant belief-forming process—on the basis of the underlying mechanisms and the conditions it operates in—and whatever the property ι that the relevant belief-forming process is supposed to generate beliefs about, the suggested approach still applies.

- True Negatives, TN, refer to the elements that the classification process identified as not bearing ι and, indeed, they do not bear ι (e.g., to those elements that they were identified as objects that are not cats and indeed they are not cats).
- False Positives, FP, refer to the elements that the test identified as bearing ι but they don't (e.g., to those elements that they were identified as cats when in fact they aren't).
- False Negatives, FN, refer to the elements that the classification process identified as not bearing ι but in fact they do (e.g., to those elements that were identified as objects that are not cats but in fact they are).

Given the above, the following also holds:

- P = TP + FN
- N = TN + FP

With the above terminology in place, we can now present the basic functions for assessing the reliability of binary classification processes, as these have been introduced by statisticians and computer scientists (indicatively, see Tharwat 2021). In addition to being clear about the following reliability indicators, it is also important to keep in mind that these indicators are distinct from one another in that a mechanism's scoring highly with respect to one of the following indicators does not entail that it will score highly with respect to any of the other indicators. Below, I offer the statistical definitions of the dimensions of reliability that the paper focuses on, without listing the reasons for which each kind of reliability is distinct from the others. Interested readers can find a detailed list of these reasons in the Appendix.

3.1. Precision

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Precision (also referred to as Positive Predictive Value, PPV) measures the likelihood that an element is an actual P (e.g., a cat), given that the mechanism classified it as such. Having high Precision means that if the relevant mechanism outputs that an element is a positive (e.g., a cat), then it is highly likely that it is a positive indeed. This can be formalised in the following way:

$$\text{High Precision} \Rightarrow \text{TP} \gg \text{FP}$$

3.2. Negative Predictive Value (NPV)

$$NPV = \frac{TN}{TN + FN}$$

NPV is the opposite of Precision. It measures the likelihood that an element is an actual N (e.g., an object that is not a cat), given that the mechanism classified it a such. Having high NPV means that if the relevant mechanism outputs that an element is a negative (e.g., an object that is not a cat), then it is highly likely that it is a negative indeed. This can be formalised in the following way:

$$\text{High NPV} \Rightarrow TN \gg FN$$

3.3. Recall

$$Recall = \frac{TP}{P} = \frac{TP}{TP + FN}$$

Recall (also referred to as Sensitivity) measures the percentage of existing positives (e.g., cats) that the mechanism will identify as such. Having high Recall means that the relevant mechanism will correctly identify most existing positives (e.g., cats).[8] This can be formalised in the following way:

$$\text{High Recall} \Rightarrow TP \gg FN$$

3.4. Specificity

$$Specificity = \frac{TN}{N} = \frac{TN}{TN + FP}$$

---

8. Recall might be reminiscent of Nozick's (1981) adherence condition: If p were true, then S would believe that p. Note, however, that Nozick's condition is formulated in modal terms. It requires that, in near-by possible worlds where p is true, if the agent employed the same belief-forming process, she would still believe that p. In contrast, Recall measures the probability that positives will be detected in the actual world—i.e., the environment and conditions in which the mechanism is, in fact, employed. This probabilistic, rather than modal, approach—concerned only with the performance of the mechanism in the actual world—is true of all the reliability indicators and corresponding dimensions of reliability that this paper explores, and it marks an important difference between the present approach and modal approaches to the notion of reliability.

Specificity is the opposite of Recall. It measures the percentage of existing negatives (e.g., objects that are not cats) that the mechanism will identify as such. Having high Specificity means that the relevant mechanism will correctly identify most existing negatives (e.g., objects that are not cats). This can be formalised in the following way:

$$\text{High Specificity} \Rightarrow \text{TN} \gg \text{FP}$$

4. Dimensions of Reliability and Epistemic Goals

The above offers precise definitions of the dimensions of reliability we are here going to focus on, and Appendix A demonstrates how they are distinct from one another. In the introductory section, we also noted that these distinct kinds of reliability may serve different epistemic goals. That is, differences in what agents aim at knowing call for high scores along different dimensions of reliability. It is also the case that certain epistemic goals may require a combination of them. A direct consequence of these points is that a shift in the agent's epistemic goal may render an otherwise reliable and knowledge-conducive process entirely inappropriate with respect to the epistemic task at hand. To get a better grip on this, it is worth considering a few examples of how certain epistemic goals may call for high levels of reliability along several dimensions.

First, take high Precision, which is required when it is important to be mostly correct about positively identified elements, but one can afford to ignore a significant percentage of positives (i.e., exhibit low Recall), does not care about frequently misidentifying positives as negatives (i.e., exhibit low NPV) and can afford to ignore a significant percentage of negatives (i.e., exhibit low Specificity). Say the epistemic agent is an art collector who wants to know whether some painting is an original Picasso in order to buy it:

EG_{AC}: When judging that an object x is a Picasso, know that x is a Picasso

With respect to this epistemic goal, all the agent needs is that the relevant belief-forming process exhibit high Precision. But a change of her epistemic goal can easily change Precision's efficiency in rendering the relevant mechanism knowledge-conducive. Say the agent wants to know every painting that is an original Picasso, because she wants to buy them *all*:

EG_{AC}': When judging that an object x is a Picasso, know that x is a Picasso; *and* know most (ideally all) Picassos

To achieve this new epistemic goal, the relevant belief-forming process would also need to exhibit high Recall. What this means is that if we focused on Precision alone to judge whether the agent's belief-forming process is reliable and thereby knowledge-conducive, we would have to conclude that it is, no matter what her epistemic goal is. Yet, when the agent wants to also know all original Picassos (i.e., when her epistemic goal is EG_{AC}', rather than EG_{AC}), high Precision alone would fail her, as she would be prone to missing out on many of them.

Now, take NPV. High NPV is appropriate when one needs to have a high success rate with respect to negatively identified elements but can afford to ignore a significant percentage of negatives (i.e., exhibit low Specificity), does not care about regularly misidentifying negatives as positives (i.e., exhibit low Precision) and can afford to ignore a significant percentage of positives (i.e., manifest low Recall). Say the epistemic agent wants to know whether incoming emails are not important in order to delete them:

> EG_E: When judging that email x is not important, know that x is not important

With respect to EG_E, all the agent needs is that the relevant belief-forming process exhibit high NPV. But a shift in her epistemic goal can easily change NPV's efficiency in rendering the relevant belief-forming process knowledge-conducive: Say that now she wants to also know all stored emails that are not important in order to delete them all (perhaps the memory of her email account is running full). In this case, her epistemic goal would be as follows:

> EG_E': When judging that email x is not important, know that x is not important; *and* know most emails that are not important.

To satisfy this new epistemic goal, the relevant mechanism would also need to exhibit high Specificity. Moreover, notice that having the relevant mechanism exhibiting only high Precision would be irrelevant to either of the above epistemic aims. Why should it matter that the agent is correct most times she thinks an email is important (i.e., manifest high Precision)? Even if Precision is high, she may still often claim that emails are not important when in fact they are (i.e., manifest low NPV), thus failing her first epistemic goal. Similarly, when manifesting high Precision, it can still be the case that a big percentage of the emails that are not important is incorrectly judged to be important (i.e., manifest low Specificity), thereby also failing her second epistemic goal.[9] So again, judging

---

9. This will be the case, if the majority of the emails she receives are important (i.e., positives). For an explanation of this point, see Section 5.

whether the agent's belief-forming process is reliable and thereby knowledge-conducive by focusing on Precision alone would deliver the wrong result.

Next is Recall. High Recall is useful when one cannot afford to ignore a significant percentage of positives, but can afford to ignore a significant percentage of negatives (i.e., exhibit low Specificity) and does not so much care about frequently misidentifying negatives as positives and *vice versa* (i.e., exhibit low Precision and NPV). Say the epistemic agent is a doctor who must know most (ideally all) patients who have contracted a virus in order to give them a cheap medicine that is amply available:

EG$_D$: Know most (ideally all) individuals with the virus

For EG$_D$ to be satisfied, it is necessary that the relevant mechanism exhibit high Recall.[10] But, again, a change in the agent's epistemic goal can easily affect the corresponding mechanism's ability to deliver knowledge. Say that the relevant medicine is in shortage, such that the doctor needs to know most individuals who have contracted the virus and needs to also know, for each positively identified individual, that they have contracted the virus indeed:

EG$_D$': Know most (ideally all) individuals with the virus; *and* when judging that individual x has the virus, know that x has the virus

In this case, the relevant mechanism must also exhibit high Precision. And yet notice that with respect to either of the doctor's epistemic goals, using Precision alone to judge whether her belief-forming process is knowledge-conducive would be insufficient. Even if the doctor is most times correct when judging one to be a patient (i.e., manifest high Precision), there can still be a lot of patients that she mistakes for healthy individuals. To avoid this, it is necessary to check whether her belief-forming process also manifests high Recall.

Finally, think of Specificity. High Specificity is fitting when one cannot afford to ignore a significant percentage of negatives, but can afford to ignore a significant percentage of positives (i.e., manifest low Recall) and does not so much care about her success rates with respect to either negatively or positively identified elements (i.e., exhibit low NPV and Precision). Say the agent is an investment consultant whom her client is relying upon to know most investments that are not profitable.

EG$_{IC}$: Know most (ideally all) investments that are not profitable.

---

10. As I discuss in Section 5, it is not clear whether high Recall is sufficient for knowing most *p*s, or whether some reasonable level of Precision (not as high as that required for 'know-that' but perhaps considerably above chance) might also be required.

With respect to this goal, it is necessary that the relevant belief-forming process exhibit high Specificity (i.e., ignore a small percentage of the investments that are not profitable).[11] But, once again, a change in the agent's epistemic goal can easily affect the corresponding mechanism's ability to deliver knowledge. Say her client is likely to be dissatisfied with being often advised against investments that are in fact profitable. In this case the consultant would also need to know, for each negatively identified investment, that it is not profitable *indeed*:

> EG_{IC}': Know most (ideally all) investments that are not profitable, *and* when judging that investment x is not profitable, know that x is not profitable.

To satisfy this new epistemic goal, the relevant mechanism would also need to exhibit high NPV. Moreover, notice, once again, that judging whether the relevant belief-forming process is knowledge-conducive on the basis of Precision would be entirely misleading. Being most times correct when judging investments to be profitable (i.e., manifest high Precision) could still expose the consultant to ignoring a big percentage of the investments that are not profitable, thereby failing her first epistemic goal.[12] Similarly, high Precision, in this case, won't prevent the consultant from often judging that investments are not profitable when in fact they are (i.e., manifesting low NPV), thereby failing her second epistemic goal.

Now, if the above is correct, there are good reasons for thinking that different epistemic goals—that is, what the agent aims at knowing—call for varying levels of reliability along different dimensions. This observation points to the need for upgrading reliabilism to *dimensional reliabilism*. According to this multidimensional approach to reliabilism, different epistemic goals call for high levels of reliability along different dimensions (and perhaps their combinations too); thus, assessing whether a process is knowledge-conducive depends on whether

---

11. As with Recall and knowing most *p*s, it is not clear whether high Specificity is sufficient for knowing most not-*p*s. It is possible that some reasonable level of NPV (again, not as high as that required for 'knowing that not-*p*', but perhaps considerably above chance) might also be required. For further discussion, see Section 5.

12. As with fn. 9, this would be the case if the majority of possible investments were profitable (i.e., positives). In such a financial situation, lacking high Specificity could be particularly dangerous. Due to the overwhelming majority of profitable investments, one could drop their guard leading to increasingly aggressive investments. At the same time, if one is unable to detect most of the unprofitable investments—however unlikely they are—could lead to a financial disaster by eventually investing a disproportionately large amount of money on the very few unprofitable investments. Indeed, it is not difficult to imagine how a con artist could take advantage of this in order to construct a financial situation wherein the investor could manifest high Precision but low Specificity in assessing the profitability of investments.

it manifests high reliability along the appropriate dimension(s), given the epistemic goal at hand.[13]

5. Correctness and Breadth of Scope: Know-That and Know-Most

The above examples in support of dimensional reliabilism are meant to indicate that different epistemic goals require varying scores of reliability across different dimensions; moreover, certain goals may require combinations of high scores across multiple dimensions. At the same time, however, the foregoing examples also reveal an interesting pattern that one may attempt to use against dimensional reliabilism: On one hand, high Precision and high NPV are most clearly associated with epistemic goals that relate to 'know-that'; on the other hand, high Recall and high Specificity seem to mostly serve epistemic goals that relate to a different kind of knowledge that is best captured by the 'know-most' and 'know-all' locutions. This observation raises two important worries for dimensional reliabilism.

First, given that mainstream epistemology has traditionally focused on 'know-that,' are Recall and Specificity of any importance to mainstream epistemology? Secondly, even if Recall and Specificity do have important roles to play, could it be the case that a belief-forming mechanism that manifests high Precision and/or NPV—that is, the two dimensions that relate to know-that'—ensures that high Recall and high Specificity are also manifested? If that's the case, then one could argue that there is a single 'master' notion of reliability—one that covers either high Precision or high NPV (or maybe both simultaneously)—which relates directly to 'know-that' and which can do all the relevant epistemic work, including the work of high Recall and high Specificity. In that case, high Recall and high Specificity would not seem as important after all and the associated 'know-most/all' would seem easily reducible to 'know-that.'

In this section, I will take up the second question concerning the reducibility of 'know-most/all' to 'know-that'. The rest of the paper will then deal with the

---

13. One may worry that Recall and Specificity are not relevant to knowledge and that the associated locutions 'know-most' and 'know-all' do not capture a genuine type of knowledge. However, as the examples of this section make apparent, 'know-most' and 'know-all' can be and often are unproblematically and rather naturally used in everyday contexts. This suggests that 'know-most' and 'know-all' are common—even if underexplored by epistemologists—epistemic states. In support of this claim, Section 6 also argues that 'know-most/all' and the dimensions of reliability that distinctively underly this kind of knowledge (i.e., Recall and Specificity) can help alleviate a common type of ignorance, to which 'knowledge' is commonly assumed to be the antidote.

first question regarding the epistemic importance of the dimensions of Recall and Specificity.

To start answering the second issue, let us return to the following pressing question: Could there be a 'master' notion of reliability, upon which—*pace* dimensional reliabilism—all assessments of reliability and knowledge may depend? Given epistemologists' long-standing focus on 'know-that' and the aforementioned pattern between the different dimensions of reliability and the kinds of epistemic goals they seem to primarily serve, a promising candidate for playing this role is what we may here refer to as 'Correctness.' Think of Correctness as the common denominator between Precision and NPV. A belief-forming process with high Precision will likely be correct with respect to positives. When it outputs that an element is a positive, it is highly likely that it is a positive indeed. Similarly, a belief-forming process with high NPV will likely be correct with respect to negatives. When it outputs that an element is a negative, it is highly likely that it is a negative indeed. In contrast, we can see that the common denominator between Recall and Specificity is what we may here refer to as 'Breadth of Scope.' A process with high Recall has high Breadth of Scope with respect to positives in that it will not ignore a big percentage of the existing positives. Similarly, a process with high Specificity has high Breadth of Scope with respect to negatives in that it will not ignore a big percentage of the existing negatives.

With the above in place, a promising way to resist the need for upgrading to dimensional reliabilism would be to claim that, given epistemologists' focus on 'know-that', what they may have so far had in mind when they thought that a process was reliable was the dimension of Correctness with respect to either positives or negatives. If we can also demonstrate that high Breadth of Scope (in the form of either high Recall or high Specificity) can be reduced to high Correctness, then we may have a strong argument against the upgrade to dimensional reliabilism.

This is not a simple matter, however. High Correctness with respect to positives (i.e., high Precision) does not entail high Breadth of Scope with respect to positives (i.e., high Recall). There might still be a significant percentage of positives that are incorrectly identified as negatives, and which are thus ignored. Similarly, high Correctness with respect to negatives (i.e., high NPV) does not entail high Breadth of Scope with respect to negatives. There might still be a significant percentage of negatives that are incorrectly identified as positives, and which are thus ignored.[14]

---

14. For similar reasons, Correctness cannot be reduced to Breadth of Scope. However, I do not discuss this option in the main text, because epistemologists seem to prioritise Correctness rather than Breadth of Scope. Discussing this possibility would therefore undermine the suggested deflationary strategy against introducing dimensional reliabilism.

Now, at this point, one might re-join by denying that epistemologists have so far focused *either* on Precision *or* on NPV. Rather, they may suggest, when epistemologists claim that a process is reliable and thereby knowledge-conducive what they mean is that it manifests high *overall* Correctness—that is, high Correctness with respect to *both* positives and negatives. This is entirely possible given that the subtle distinction between Precision and NPV has so far gone unnoticed. But then, if we grant this assumption, we have to revisit the reducibility question and, this time, ask: Can high overall Correctness ensure high overall Breadth of Scope (i.e., high Breadth of Scope with respect to positives and high Breadth of Scope with respect to negatives)?

As it happens, the answer is again negative. Take a mechanism that exhibits both high Precision and high NPV. Even if the mechanism does not often misidentify negatives as positives (due to high Precision), and even if the mechanism does not often misidentify positives as negatives (due to high NPV), it could still be the case that either Recall or Specificity are low. In cases where the number of positives in the population is much higher than the number of negatives, the number of negatives that can be correctly identified (*i.e.,* the number of TN) will be close to the number of FP (despite the fact that, due to high Precision, the mechanism very rarely misidentifies negatives as positives).[15] Consequently, a large percentage of negatives will be ignored (i.e., Specificity will be low). Conversely, in cases where the number of negatives in the population is much higher than the number of positives, the number of positives that can be correctly identified (*i.e.,* the number of TP) will be close to the number of FN (despite the fact that the mechanism, due to high NPV, very rarely misidentifies positives as negatives).[16] Consequently, a large percentage of positives will be ignored (i.e., Recall will be low).[17]

Therefore, it does not seem possible to reduce high overall Breadth of Scope (i.e., both high Recall *and* high Specificity) to high overall Correctness (i.e., both high Precision *and* high NPV). This indicates that the largely ignored locutions of 'know-most/all' may not be reducible to 'know-that'—at least not from a reliabilist perspective. The reason is that 'know most/all' require that the underlying belief-forming mechanism manifests high scores across certain kinds of reliabil-

---

15. When the number of positives is very high, high Precision (TP ≫ FP) does not mean that the actual number of FP will be low, especially when compared to N and TN.

16. When the number of negatives is very high, high NPV (TN ≫ FN) does not mean that the actual number of FN will be low, especially when compared to P and TP.

17. As a worked example, consider a population with 18 positives of which 9 are correctly identified, and 82 negatives of which 81 are correctly identified. In this scenario, Precision = 0.9, NVP = 0.9, Specificity ≈ 0.99, but Recall = 0.5. Conversely, if P = 82, TP = 81, N = 18 and TN = 9, then Precision = 0.9, NPV = 0.9, Recall ≈ 0.99, but Specificity = 0.5. In general, there is no clear-cut way to draw direct connections between any of these kinds of reliability, because, as the above indicates, their values depend on each other as well as the corresponding value of the population's prevalence (i.e., the percentage of positives in the population = P / (P + N)).

ity that cannot be ensured by high scores in the kinds of reliability required for 'know-that.' Put another way, 'know-most/all' seems irreducible to 'know-that', because having both the requisite justification for knowing that p and the requisite justification for knowing that not-p does not entail that one possesses the justification required for knowing most (ideally all) ps and not-ps.

But if 'know-most/all' does not reduce to 'know-that,' could it at least be the case that 'know-that' is a necessary component of 'know-most/all'? It may be pointed out, for example, that if one were in a position to detect most ps and *not-p*s, but manifested low Correctness by manifesting low Precision and/or low NPV, it would be inappropriate to claim that they nevertheless *know* (just because of high Recall and Specificity) most ps and *not-p*s. In other words, it may be pointed out, low Precision and low NPV are incompatible with 'know-most/all.'

Consider for example the cat-identifying mechanism of Section 2 and a population consisting of 1090 animals that are not cats—of which 1000 are correctly identified as not-cats (i.e., TN=1000 and FP=90)—as well as of 11 cats—of which 10 are correctly identified as cats (i.e., TP=10 and FN=1). In this case, Recall ≈ 0.91, Specificity ≈ 0.92, NPV ≈ 0.99, but Precision = 0.1. Despite the high Recall of this mechanism, one could object that it is not possible to know-most cats on the basis of this mechanism, because, on its basis, when one judges that an animal is a cat one cannot *know that* the target animal is a cat—claiming that something is a cat, after all, turns out to be correct only one out of ten times (i.e., Precision = 0.1). In other words, the worry may go, such a mechanism cannot be said to know most cats because there seems to be something wrong with the 'idea' of cats it is operating on—there simply are too many animals that it mistakes for cats. And if one can be so easily mistaken about what a cat is, they cannot really *know* most cats, even though, were they interested in obtaining as many cats as possible, due to promiscuous ordering, they would miss out on very few of them. Put simply, judging to be a cat anything that remotely looks like a cat, so that, in the end, very few cats will be ignored, prevents one from knowing most cats: Knowing most cats, the suggestion is, does not only require that one miss out on very few cats—in doing so, one must also be in a reasonably good position telling cats from cat-looking animals apart.

Admittedly, the above suggestion is not without merit.[18] But even if we accept it, it does not directly lead to the conclusion that 'know-that' is necessary for 'know-most/all.' Knowing that p requires a sufficiently high level of Precision—presumably, most people would consider 0.9 as sufficient and anything below

---

18. Though in the remainder of the discussion I accept it, I am not entirely clear I agree with it. One of the reasons I worry about it is that long exposure to mainstream epistemology (which is primarily focused on 'know-that') may give epistemologists the impression that all kinds of knowledge must closely resemble knowledge-that—even if this is not true. Experimental philosophy could help settle this issue.

0.8 would likely be unacceptable, though, depending on the process, different levels might be deemed acceptable each time. It is not clear, however, that the same level of Precision is required for knowing most *p*s. Say one is in a position to miss out on very few cats by manifesting very high Recall, but the precision of their cat-identifying mechanism, while well above chance, falls short of the reliability required for knowing that something is a cat—say it is only 0.7. Or say that one can list most Michelin-starred restaurants in their city, but only seven out of the ten restaurants that they list are indeed Michelin-starred. Wouldn't in these cases still be appropriate to claim that the agent knows most cats or that they know most Michelin-starred restaurants, even though the moderate Precision of the underlying mechanisms would not support them in knowing that something is a cat or knowing that a restaurant is indeed Michelin-starred? Intuitively, the answer here seems to be positive; claiming, in such cases, that agents 'know-most', even though they do not quite 'know-that,' sounds fair and unproblematic. If that's correct, however, then not only is it not possible to reduce 'know-most/all' to 'know-that' but 'know-that' may not even be necessary for 'know-most.' Rather, 'knowing most/all' may only require a reasonable level of Correctness (perhaps, considerably above chance).

In any case, whatever the answer regarding the necessity of 'know-that' for 'know most/all' may be, the above indicates that there are good reasons for thinking that 'know-most/all' cannot be dealt with merely by appealing to ''know-that.' This leaves us with the less technical worry that 'know-most/all' and Breadth of Scope might not be epistemically important—at least not important enough to warrant the attention of mainstream epistemology. Is this true?

6. Correctness, Breadth of Scope, Knowledge and Ignorance

As we have seen, it does not seem possible to do away with Breadth of Scope and 'know-most/all' by reducing them to Correctness and 'know-that.' Yet there might be another way to resist the shift to dimensional reliabilism: Instead of seeking to reduce Breadth of Scope to Correctness, one could argue that, from a reliabilist perspective, Breadth of Scope is unrelated to knowledge assessments. Arguing in this direction would clearly reinforce the mainstream epistemological *status-quo*, whereby the focus has been primarily on 'know-that.' But is it correct to claim that Breadth of Scope is irrelevant to knowledge assessments? After all, as we have already seen, on many occasions, knowing most *p*s or not-*p*s is at least as important as knowing that *p* or not-*p*, and on other occasions it is crucial to know both *that p* or not-*p* and *most p*s or not-*p*s. So how might one deny that Breadth of Scope is relevant to knowledge assessments such that reliabilists, *qua* reliabilists at least, should not be concerned with it?

To see how such a strategy might go, consider Goldman's (1986: 122) notion of 'power', which sounds similar to Breadth of Scope, though, as I will shortly explain, it is not quite the same:

> Reliable cognitive mechanisms are an antidote to error, but not necessarily an antidote to ignorance. Reliable processes guarantee a high truth ratio among the beliefs they generate. But they may generate very few beliefs indeed; they may leave the cognizer with very little information, even on the issues that interest him or her the most. The antidote to ignorance is not reliability, but (intellectual) power. Powerful cognitive mechanisms are (roughly) mechanisms capable of getting a relatively large number of truths. An intelligent cognitive system, it seems clear, is not simply a reliable system.

Goldman's power and Breadth of Scope are similar in that they are both concerned with alleviating 'ignorance.' Nevertheless, we should be clear that the two epistemic notions are by no means the same. As Goldman (1986: 123) further notes: "Power [is] a function of the proportion of questions it [a cognitive system] wants to answer that it can answer (correctly), or the proportion of problems undertaken that it can solve correctly." So, on one hand, Goldman's power refers to a system's ability to address a big proportion of the questions and problems it is interested in. Breadth of Scope, on the other hand, refers to the ability of a process to detect a high percentage of the positives or negatives it is looking for. Thus, the two notions seek to alleviate ignorance in different senses of the term. On one hand, power is meant to prevent *cognitive systems* from being reliable but highly specialised in the sense of being capable of generating correct beliefs in a few narrow domains, while failing to generate (thus ignoring) true beliefs in other important domains. On the other hand, Breadth of Scope is meant to prevent *belief-forming processes* from failing to detect (and thus ignore in another sense of the term) a significant percentage of the items they are meant to be looking for.

Despite the fact that the two notions are not the same, I here mention Goldman's discussion of power because I am interested in his strategy for accommodating it within epistemology. Notably, Goldman does not associate power with knowledge but with the distinct notion of 'intelligence.' Presumably, the reason for this move is Goldman's programmatic commitment to analysing knowledge and justification in terms of reliability. Since Goldman does not have a way to account for power in terms of reliability (i.e., in terms of a high preponderance of true over false beliefs), he avoids claiming that power is directly relevant to knowledge. So, the question we need to address here is whether one could deny that Breadth of Scope is directly relevant to knowledge in a similar manner.

In response, the preceding indicates that Breadth of Scope can be readily captured in terms of the standard definition of reliability as a 'high preponderance of true over false beliefs.' From reliabilism's perspective, it would therefore seem arbitrary to discount Breadth of Scope as the kind of reliability that can be directly relevant to knowledge—especially when, as we have already demonstrated, high Breadth of Scope (i.e., high Recall and/or Specificity) is at least necessary for obtaining knowledge in a number of different cases (think of EG_{AC}', EG_E', EG_D, EG_D', EG_{IC}, EG_{IC}' as discussed in Section 4). Overall, then, accounting for *knowing most p*s (or not-*p*s) can be as important as accounting for *knowing that p* (or not-*p*), and reliabilism (or, more precisely, its upgraded version of *dimensional* reliabilism) is now in a position to analyse and assess attempts to satisfy both kinds of epistemic goals (and their combinations).

What is more, expanding reliabilism's purview in this way should be a welcome result from a folk-psychological perspective. For, while Goldman might be correct in claiming that the antidote to the kind of ignorance he has in mind is 'intellectual power,' it is also widely assumed that the opposite of ignorance is *knowledge*. So, being in a position to explain how reliability in the form of high Breadth of Scope is crucial to knowing a sufficiently big percentage of what one is looking for is a promising way for spelling out, in reliabilist terms, this (at least) equally common assumption: That the way to alleviate (a common kind of) ignorance is through having knowledge—the kind of knowledge that is best captured by the locutions 'know-most/know-all', and which necessarily requires that the underlying belief-forming mechanisms manifest high Breadth of Scope.[19]

---

19. There is an additional point worth considering here, regarding the somewhat related issue of suspending belief: One may worry that, on the suggested approach, the mechanism being assessed for reliability will either output a judgment of the form, "X is an F", or a judgment of the form, "X is not an F". But human believers are not always opinionated in this way: we often suspend judgment on matters. This worry can be easily handled, by noting that even though the mechanism may only output either that "X is F" or that "X is not an F", this does not necessarily mean that the bearer of the mechanism will end up believing the generated output. For example, if the agent has undercutting defeaters against the relevant reliability dimension of her mechanism, she may well suspend judgment by refraining from believing the delivered output. Another version of this objection would be to claim that not all judgments are binary. Something might be neither good nor bad. Similarly, an object may be neither beautiful nor ugly. Accordingly, dimensional reliabilism cannot account for non-binary belief-forming processes. On a first pass, one may claim that most non-binary judgments are value judgments instead of epistemic judgments. Nevertheless, if even if that's incorrect, such that several non-binary epistemic judgments exist (which they likely do), dimensional reliabilism would still be of significant import, as it would apply to many—even if not all—epistemic belief-forming processes. As noted in the beginning of Section 3, binary classification seems applicable to most of our basic epistemic belief-forming processes.

7. The Evolution of Reliability and Its Different Dimensions

According to evolutionary biology all species have evolved to exhibit traits that are conducive to coping within their natural environments. Instead of being the result of some intelligent design, these traits have, in brief, evolved by means of a long process of natural selection, which extends far in the past of the evolutionary history of each species. Likewise, according to evolutionary epistemology, organisms' epistemic capacities are the products of the same evolutionary process (see also Bradie and Harms 2023).

Conceiving of epistemic cognitive processes as the product of natural selection provides evolutionary epistemologists with a simple yet plausible argument for the reliability of naturally evolved belief-forming processes. Briefly, if a belief-forming process is not reliable, then any organisms bearing the relevant process will be unlikely to survive until the age of reproduction. In result, the unreliable process will not be passed on to later generations. Conversely, any belief-forming process that has passed the test of natural selection is a process that is in fact reliable.

This is an important claim for evolutionary epistemology and especially reliabilist approaches to it, such as proper functionalism. For example, in defending this claim, Peter Graham notes: "This matters a great deal to me, as I believe perceptual epistemic warrant is constitutively associated with perception having the etiological function of inducing reliably true perceptual beliefs, where biological functions are a species of etiological function" (2014a: 2).[20] Nevertheless, the tight connection between evolution and epistemic reliability has also been the target of much criticism over the years, with Stich (1985), Kitcher (1992) and Pernu (2009) all raising objections against it (or close formulations of it). While it is beyond the scope of this paper to fully defend the claim that evolved belief-forming processes are reliable against all existing objections, it will be useful to demonstrate how dimensional reliabilism can come to its rescue with regards to what has come to count as a pressing counterexample.

---

20. As Graham further notes (2014a: fn. 2), on his view, "warrant consists in normal functioning when the belief-forming process has forming true beliefs reliably as an etiological function. Natural selection is sufficient, but not necessary, for etiological functions." An etiological function, on Graham's view, is a function "that exists or continues to exist in terms of its consequences, because of a feedback mechanism that takes consequences as input and causes or sustains the item as output" (Graham 2014b: 18). So, on this version of proper functionalism, when a perceptual mechanism has evolved by natural selection, it will have the etiological function of reliably producing true beliefs; thus, when it functions normally, it will be reliable. Overall, then, Graham holds that biological perceptual mechanisms are (when they function normally) reliable, because they have evolved via natural selection to do just that; and that if a perceptual mechanism has evolved via natural selection, then it is reliable (provided also that it functions normally).

The counterexample I have in mind was first introduced by Tyler Burge (2010) and later discussed by Graham (2014a). Burge considers the case of rabbits and their capacity to accurately represent danger. The worry is that rabbits appear to be easily frightened in that they often behave as though danger is present when in fact it is absent. In his attempt to deal with this potentially worrying counterexample, Graham (2014a: 22) writes: "such [danger-detection] mechanisms are often unreliable, for false positives ('danger is present' when there is nothing to fear) outnumber true positives ('danger is present' when it's time to run)." Accordingly, there appears to be a naturally selected belief-forming process, which is unreliable. Nevertheless, even if unreliable, Graham notes that the rabbit danger-detection mechanism is *effective*. Specifically, he writes (2014a: 26):

> Many predator detectors work like this, where the representation of danger is not very reliable; it often represents the presence of danger when there is nothing to fear. Though effective—they "fire" almost every time danger is present and so keep the organism safe from harm, or at least give the animal a fighting chance—they frequently fire when danger isn't present, and so are not very reliable. Nature has settled on such a way of avoiding predators because false negatives ("there is no danger present; I'm safe" when danger is lurking) are so much worse than false positives ("danger is present, run!" when there's nothing to fear). If the animal overestimates the chances of danger and runs away at the slightest sign, it will effectively avoid predators when they are present, even if it frequently runs away when, in fact, it is perfectly safe. [. . .] Nature settled on an unreliable but effective device.

So, let us ask: Is Graham correct in thinking that such a danger-detecting mechanism is unreliable? And what does he mean when he says that it is, at least, *effective*?

Graham judges rabbits' ability to detect danger as being unreliable, because "false positives ('danger is present' when there is nothing to fear) outnumber true positives ('danger is present' when it's time to run)." Given the terminology introduced in Section 3, it seems clear that Graham refers to Precision. This shouldn't really come as a surprise, given the preceding points that high Precision is clearly required for 'know-that'. Nevertheless, as Graham further notes, such danger-detection mechanisms, though they fail to exhibit high Precision, they are at least *effective*. "They 'fire' almost every time danger is present and so keep the organism safe from harm, or at least give the animal a fighting chance."

So, given these points, Graham (2014a: 26) concludes his discussion of Burge's counterexample, in the following way:

> Nature settled on an unreliable but effective device, effective because accurate often enough. Most of our perceptual states and systems, however, are not like this. Most are reliable, and contribute to fitness by being reliable. Unreliable danger detectors are the exception that, so to speak, proves the rule.

Now, it is doubtful that proponents of evolutionary reliabilist epistemology would find this response particularly appealing—admitting, after all, there is an exception to the rule (that the evolution of epistemic mechanisms ensures their reliability) is to admit that the rule is potentially wrong. But does Graham need to concede this much? Is it actually correct to admit that the mechanism that keeps rabbits safe is unreliable on the grounds that it manifests low Precision?

I want to suggest that dimensional reliabilism may provide Graham with the means to offer a better response to the 'rabbits counterexample.' To see how, we only need to pay attention to what Graham means when he claims the mechanism is *effective.* If we invoke the terminology of Section 3, we will come to see that Graham's points essentially amount to claiming that the relevant mechanism exhibits high Recall. That is, when Graham points out that rabbits' danger-detection mechanisms are at least 'effective', he is effectively making a claim that is far more helpful to his view (and to evolutionary epistemology more generally) than it may initially sound. From the point of view of dimensional reliabilism, he is, essentially, making the claim that rabbits' danger detection mechanisms *are* reliable, because they exhibit high Recall.

This may sound problematic if it is assumed that Precision is all there is to reliability. But from an evolutionary perspective, focusing, in this case, on Recall rather than Precision makes sense. Rabbits are vulnerable creatures that need to stay safe. So long as they are reliable by exhibiting high Recall with respect to danger, they can live a long and happy life—even if they often misrepresent safe situations as dangerous ones. Put another way, given their vulnerable nature, it makes sense that they have evolved a danger-detection mechanism that is epistemically reliable in such a way that it allows them to be 'better safe than sorry' (see also Stich 1985: 125).

The present approach therefore can be used to defend the claim that evolved belief-forming processes are reliable against the rabbit counterexample. It demonstrates that even though rabbits' danger-detection mechanism manifests low Precision, this is not sufficient grounds for thinking that the relevant mechanism is overall unreliable. Given rabbits' *epistemic goal*—that is, detect most (ideally all) situations that are dangerous—Precision is not the right dimension of reliability to prioritise when assessing whether their danger-detection mechanism is reliable.

Being most often correct when judging a situation to be dangerous but being unable to detect *most* dangerous situations would be disastrous for them. Instead, rabbit's danger-detection mechanism needs to manifest high Recall—which in the case of rabbits is remarkably high. Thus, given their epistemic goal, rabbits' danger detection mechanism *is* reliable after all and, contrary to Graham's concession, they do not represent an exception to the rule that evolution ensures epistemic reliability.[21]

Of course, it may be further pointed out that the debate on evolutionary epistemology may have been framed, for the most part, in terms of a general notion of reliability or accurate representation, but, really, participants to the debate have been interested in the kind of reliability that can ensure *knowledge*. This suggestion could explain, for example, why the debate seems to have so far focused on Precision—the narrative should be familiar by now: Mainstream epistemologists have been primarily interested in 'know-that', and high Precision, as we have seen, is intricately related to 'know-that.' This may well be true, though it hardly justifies the debate's exclusive focus on Precision. As it has been here argued, it is questionable whether Precision, or, more broadly Correctness, and the associated 'know-that' are all that should matter epistemically. 'Know-most/all' appears to be an equally important epistemic goal and the debate on evolutionary epistemology should pay equal attention to it.

And yet, it may be further argued, even if we grant that the debate on evolutionary epistemology should not exclusively focus on 'know-that' and high Precision, but should, instead, also focus on 'know-most/all' and Breadth of Scope, Precision may still play a role in assessing whether evolution can ensure the kind of reliability required for 'know-most.' As noted toward the end of Section 5, in addition to high Breadth of Scope, some degree of Correctness *may* also be necessary for 'know-most,' just not to the same level required for 'know-that'—perhaps 'know-most' requires that Precision be only considerably above chance. With that in mind, the debate on evolutionary epistemology would then need to revisit the case of rabbits and ask: Can evolution ensure the kind of multidimensional reliability required for 'know-most'? In other words, while we may grant

---

21. One may react to the whole discussion regarding the case of rabbits and, specifically, its pertinence to evolutionary epistemology, by pointing out that rabbits do not represent danger; rabbits—the worry may go—do not represent any concepts at all. So, the relevant mechanism is not an epistemic mechanism but a practical one. While this may exclude rabbits' danger detection mechanism as a case study regarding the reliability of epistemic mechanisms (and their evolution), there are several human epistemic mechanisms that are similar in that they exhibit high Recall but not very high Precision. Indeed, one may focus on the danger-detection mechanism of human beings, which is difficult to deny that it represents danger. As with rabbits, in many situations, humans' danger-detection mechanism exhibits high Recall instead of high Precision. Thus, any of the above claims with respect to rabbits' danger-detection mechanism may apply equally well with respect to humans' corresponding mechanism, which is undoubtfully an *epistemic* mechanism. The focus, here, has been on rabbits rather than humans, only because of the existing literature's focus on the case of rabbits.

that rabbits are able to *detect* most instances of danger (due to the high Recall of their mechanism), are they also in a position to *know* most instances of danger? The answer to this question will largely depend (a) on how much higher than chance (if that) we need Precision to be for claiming that a mechanism can deliver 'know-most' and (b) on the answer to the further empirical question regarding the extent to which rabbits actually misrepresent danger.[22] While their Precision might be insufficient for *knowing that* danger is present, perhaps it would not be all that surprising to find out that it is good enough for *knowing* (as opposed to merely detecting) *most* instances of danger. And should this turn out to be correct, then the case of rabbits would not only fail to count as a counterexample against the claim that evolution ensures epistemic reliability; it would also fail to raise doubts against the stronger claim that evolution ensures reliability that is sufficient for (at least an important kind of) knowledge.

At any rate, the above should make clear that dimensional reliabilism can help us retain the general, causal link between evolution and epistemic reliability on the face of the rabbits example; what is more, introducing the suggested additional dimensions of reliability can add significant nuance to our understanding of the claims put forward by evolutionary epistemology. More generally, and relatedly, the preceding also highlights a broader point that has already been introduced in previous sections: Employing a mono-dimensional approach to epistemic reliability, whereby one is more likely to prioritise Precision (or, more broadly, Correctness) over other dimensions of reliability is misguided. Given epistemologists' keen interest in 'know-that,' it is hardly surprising that this may have been the case (at least implicitly) so far. But as the foregoing suggests, mainstream epistemology may need—and now has the means—to get past this. To do so, it needs to stay aware of the fact that different epistemic goals—that is, what the agent aims at knowing—seem to call for varying levels of reliability across different dimensions. This is the core message of dimensional reliabilism.

8. Applications

So far, we have seen that mainstream epistemology's implicit assumption that reliability is a mono-dimensional concept has been called into question. Functions for assessing the reliability of classification processes reveal that mainstream epistemology's definition of reliability is open to at least four different

22. Graham claims that false positives outnumber true positives, but it is not entirely clear whether this is correct or the extent to which it is correct: Measuring the Precision of rabbits' danger-detection mechanism is an open empirical question that needs to be carefully dealt with before reaching any conclusions (for instance, we need to carefully figure out what might count as a true positive and then measure rabbits' behaviour in circumstances that resemble their natural environment as much as possible).

interpretations: Precision, Negative Predictive Value (NPV), Recall and Specificity. It has also become evident that, depending on the agent's epistemic goals, these dimensions of reliability and their combinations can all play a direct role in assessing whether a process is appropriately reliable and thereby knowledge-conducive. This observation invited us to upgrade reliabilism to dimensional reliabilism, according to which, assessing whether a process is knowledge-conducive depends on whether it manifests sufficiently high reliability along the appropriate dimension(s), given the epistemic goal at hand. As a final attempt to motivate the upgrade to dimensional reliabilism, this section highlights a few ramifications of the view, across both the theoretical and the practical domain.

8.1. Theoretical Ramifications

In the previous section, it became apparent that dimensional reliabilism can have a significant input to the discussion surrounding evolutionary epistemologists' claim that evolved belief-forming processes are reliable. An additional way the view can impact mainstream epistemology concerns the interpretation of thought experiments. Consider for example the much-discussed case of Barney:

> *Barney*
> Barney is driving through the country and happens to look out of the window into a field. In doing so, he gets to have a good look at a barn-shaped object, whereupon he forms the belief that there is a barn in the field. This belief is true, since what he is looking at is really a barn. Unbeknownst to Barney, however, he is presently in 'barn-façade country' where every object that looks like a barn is a convincing fake. Had Barney looked at one of the fake barns, then he would not have noticed the difference. Quite by chance, however, Barney just happened
> to look at the one real barn in the vicinity. (Pritchard 2009: 12)[23]

Is Barney's belief-forming process appropriately reliable; and, relatedly, does Barney have knowledge in this case? The answer depends on Barney's epistemic goal—what he aims at knowing. If Barney needs to know whether the object he is looking at is a real barn, then, intuitively, he lacks knowledge. Reliabilists may account for this intuition by claiming that Barney's belief-forming process is unreliable—period. But given dimensional reliabilism this is a misleadingly crude assessment of what is going on. Granted, in this environment, Barney's barn-detection mechanism does not manifest high Precision, which is necessary for knowing

---

23. The Barney case is described in Goldman (1976) and credited to Carl Ginet.

that the object he is looking at is a real barn. Nevertheless, Barney's barn-detection mechanism still exhibits high Recall in that it can allow Barney to detect most real barns. Moreover, assuming, quite plausibly, that the countryside also contains considerably more objects that are neither barns nor barns façades, Barney's barn-detection mechanism also exhibits high NPV and high Specificity. Thus, given Barney's scores across the different dimensions of reliability, Barney does not know whether any given Barn is indeed a barn—his Precision is too low; and provided that some reasonable level of Precision (perhaps above chance) might be necessary for knowing-most/all barns, Barney may also fail to know most/all barns in the territory. Nevertheless, if Barney wants to know whether vehicles, trees, houses, cats and the like are not barns, then the high NPV of his barn-detection mechanism allows him to know such propositions (and this would be so, even if, for some strange reason, Barney had no vehicle-, tree-, house- or cat-detection mechanisms). Similarly, if Barney wants to know most/all objects that are not barns, the high Specificity of his barn-detection mechanism allows him to know that too.

The upshot is that while the low Precision of Barney's barn-detection mechanism disallows Barney from *knowing that* the barn he is looking at is a real barn and perhaps also prevents him from knowing all real barns, it would be a mistake to claim that his belief-forming process is overall unreliable and thus incapable of generating *any* knowledge. Barney's barn-detection mechanism is reliable in that it manifests high Recall, NPV and Specificity, allowing Barney to *detect most/all* barns, to *know that* many objects are not barns and to *know most/all* objects that are not barns, respectively.

Specifying, thus, the epistemic goals to which the different dimensions of reliability are relevant and the level of reliability that agents' belief-forming processes manifest along each dimension is crucial for assessing whether and what the subjects of epistemological thought experiments know or are epistemically capable of. Since much of mainstream epistemology and the debates surrounding reliabilist approaches to knowledge and justification rely on insights elicited on the basis of such thought experiments, it would be a mistake to undermine the relevance of dimensional reliabilism for their interpretation.

8.2. Practical Ramifications

Additionally, dimensional reliabilism can play a significant role in explaining reliability assessments in real life.

Potentially, there is an endless list of examples, including, to mention a rather familiar one, the way we deem friends and relatives reliable informants for, say, movie and restaurant suggestions. Most friends are reliable by being highly Precise. They rarely get it wrong. But on rare occasions, there are excellent infor-

mants who also exhibit high Recall by providing long lists of good suggestions. These are the ones we actively seek the advice of.

Another noteworthy way that dimensional reliabilism can become handy is by offering reliabilists a more nuanced and detailed way for analysing and understanding the reliability of medical tests, such as SARS-CoV-2 PCR and rapid antigen tests. As is commonly known, PCR tests are often supposed to be more reliable than SARS-CoV-2 rapid antigen tests. But what exactly does this mean? After all, SARS-CoV-2 rapid antigen tests have been found to have high or moderately high Precision, particularly high NPV and particularly high Specificity on populations without COVID-19 symptoms; a likely answer is that their Recall of asymptomatic positive individuals can be considerably low—for example, it has been reported to be 40.0% (García-Fiñana et al. 2021) and 71.43% (Fernandez-Montero et. al 2021) for two different tests.[24] This means, roughly, that of all asymptomatic subjects who were actually positive and who took, as part of these studies, one of the two rapid antigen tests, only 40% and 71.43%, in each study respectively, were correctly identified by the rapid antigen tests as positive, while the rest of the positive, tested subjects received false negative results, which could have led them into incorrectly believing that they were negative; and crucially, this was so, even though the Precision and NPV of these tests was at least moderately high: that is, respectively, when the subjects in these studies took these tests and received positive results, they were substantially likely to be positive; and when subjects received negative results, they were highly likely to be negative indeed.

This illustrates that if we only focused on Correctness in the form of either Precision or NPV, or even if we focused on overall Correctness (i.e., both Precision and NPV) we would be unable to explain the crucial sense and the full extent to which the reliability of the tested rapid antigen tests differred to that of PCRs—the tested rapid antigen tests' Correctness, after all, ranged between moderately and particularly high (depending on the test and the kind of correctness under question). By employing dimensional reliabilism, however, we can easily make sense of the situation. When it comes to testing for highly infectious viruses, such as SARS-CoV-2, we do not only need to *know that* a certain subject has or does not have the virus; there are circumstances in which we also need to *know most* (ideally all) tested subjects who have the virus. While the levels of Precision and NPV of

---

24. The study by García-Fiñana et al. (2021) focused on the reliability of the Innova LFT test; the study by Fernandez-Montero et. al (2021) focused on the Roche Antigen Test. The values for Precision, NPV and Specificity on populations without COVID-19 symptoms as reported by these studies were: Precision=90.3%, NPV=99.2%, Specificity=99.9% (García-Fiñana et al. 2021); and Precision=81.4%, NPV=99.44%, Specificity=99.68% (Fernandez-Montero et. Al 2021). For more details and nuanced analysis of the results, readers should consult the original studies. No values for any of these dimensions of reliability exist for PCR tests because PCR tests are the *gold standard* against which the reliability of rapid antigen tests is tested.

the tested rapid antigen tests may support the satisfaction of the first kind of epistemic goal (though we should also keep in mind that know-that is often thought to require more than just sufficiently high reliability), it appears that manifesting low Recall of asymptomatic positive individuals prevents them from supporting the second. On the contrary, PCRs can help in both regards.[25]

Finally, a further example of the utility of dimensional reliabilism concerns the pressing societal issue of the reliability of media and online information sources. The increasingly prevalent role that unregulated blogs and social media play in information sharing has led to the rise of fake news and alternative facts and has amplified the threat of propaganda both online and across more traditional information channels. In response, mainstream epistemology has rightly turned its focus to assessing the reliability of media (e.g., Tollefsen 2009; Coady 2012; Frost-Arnold 2014). In line with this epistemological critique of modern information resources, dimensional reliabilism offers a unique toolkit for assessing the reliability of media.

Media that regularly present false information as facts (i.e., fake news) exhibit low Precision and should, no doubt, be criticised for being thus unreliable. But would it be enough for a medium to exhibit high Precision in order to qualify as reliable? In a time where fake news and alternative facts are so widespread, is it not also desirable—perhaps necessary—that reliable media also inform their consumers about any existing information that is fake? Wouldn't a more reliable medium need to also exhibit high NPV by regularly and accurately exposing fake news? Or think about propaganda based on selective or biased reporting. A fully reliable medium cannot be merely one that manifests high Precision and high NPV. For it would be possible to score high in these reliability indicators while failing to report most relevant facts and expose most fake news. Avoiding propaganda

---

25. An anonymous referee has suggested that, contrary to the present suggestion, epistemologists should not (and presumably do not) focus on whether there is a preponderance of a certain type of true beliefs (such as true positives or true negatives) over another type of false beliefs (such as false negatives or false positives). Rather, epistemologists should be interested in whether, of all the beliefs outputted by a certain belief-forming process, there will be more true beliefs than false beliefs. In response, we can here point out that should this be the case, then, again, we would be unable to explain the sense in which the tested SARS-CoV-2 rapid antigen tests appear to be in an important sense less reliable when compared to PCRs. While the overall truth to false ratio of PCRs is better than that of the rapid antigen tests, in both cases it is considerably high, such that comparing the two kinds of tests in this manner cannot account for the qualitative (as opposed to merely quantitative) difference between them. To explain what the crucial qualitative difference is, we need to focus on the significant difference in their respective Recall. Of course, it should be noted that the foregoing constitutes a rough philosophical (as opposed to detailed medical or epidimiological) analysis in the context of a broader philosophical topic. Anyone who is interested in acquiring a detailed and thorough understanding of the reliability of the above rapid antigen tests for SARS-CoV-2 should consult the original studies, and individuals who need to be tested for SARS-CoV-2 should follow the advice of medical experts and the official authorities about what tests to use and how to proceed, depending on their individual circumstances and results.

on the basis of selective or biased reporting can only be achieved by media that also exhibit high Recall and high Specificity. They must publicise most (if not all) relevant facts and reveal most (if not all) false information (irrespectively of where this information originates from or the agenda it may serve). Only then can it be ensured that an information channel does not contribute to propaganda.

Dimensional reliabilism thus offers the resources for addressing several theoretical considerations and practical concerns surrounding the concept of epistemic reliability. While the above points are an incomplete list of the several ways in which this multidimensional approach to epistemic reliability can impact mainstream epistemology and society at large, they clearly indicate that dimensional reliabilism has far reaching ramifications on a broad range of topics and that it can therefore significantly expand the scope and reach of reliabilism.

9. Conclusion

By reference to functions for assessing the reliability of binary classification processes, the paper has argued that the popular epistemological approach of reliabilism stands to significantly benefit from the upgrade to what I have called dimensional reliabilism. Some of the theoretical advantages of this move involve bringing to the fore the importance of knowledge captured by the 'know-most/all' locutions, which, from a reliabilist perspective, do not seem reducible to 'know-that.' Moreover, the paper has demonstrated that by paying attention to 'know-most/all' and the underlying reliability dimensions of Recall and Specificity, we can defend evolutionary reliabilist epistemology from the 'rabbits counterexample' as well as provide the means for more nuanced analyses of several theoretical matters, including questions within evolutionary epistemology and the results of epistemological thought experiments. Additionally, and perhaps most importantly, on the practical side, dimensional reliabilism offers epistemologists the tools for analysing and better understanding the reliability of various information resources—from medical tests and individual informants to mass media and online information sources—by drawing important comparisons and distinctions that are inaccessible to a mono-dimensional approach to the notion of reliability. The upshot is that dimensional reliabilism has the potential to significantly expand the purview of reliabilism.

Acknowledgements

Cardiff Research Seminar for their comments on previous versions of the paper, as well as to two anonymous referees for *Ergo*, whose feedback significantly shaped the final draft of the paper. I am deeply grateful to Konstantinos Koumatos (Mathematics, University of Sussex) for his support and patience in helping me better understand the mathematical functions underlying the present arguments (any mistakes would, of course, be my own).

References

Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.

Bradie, M. and W. Harms (2023). Evolutionary Epistemology. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition). https://plato.stanford.edu/archives/spr2023/entries/epistemology-evolutionary/

Coady, D. (2012). *What to Believe Now: Applying Epistemology to Contemporary Issues*. John Wiley & Sons.

Conee, Earl and Richard Feldman (1998). The Generality Problem for Reliabilism. *Philosophical Studies*, *89*(1), 1–29.

Feldman, R. (1985). Reliability and Justification. *The Monist*, *68*(2), 159–74.

Fernandez-Montero, A., J. Argemi, J. A. Rodríguez, A. H. Ariño, and L. Moreno-Galarraga (2021). Validation of a Rapid Antigen Test as a Screening Tool for SARS-CoV-2 Infection in Asymptomatic Populations. Sensitivity, Specificity and Predictive Values. *EClinicalMedicine*, *37*, 100954.

Frost-Arnold, K. (2014). Trustworthiness and Truth: The Epistemic Pitfalls of Internet Accountability. *Episteme*, *11*(1), 63–81.

García-Fiñana, M., D. Hughes, C. Cheyne, G. Burnside, M. Stockbridge, T. Fowler, V. Fowler, M. Wilcox, M. Semple, and I. Buchan (2021). Performance of the Innova SARS-CoV-2 Antigen Rapid Lateral Flow Test in the Liverpool Asymptomatic Testing Pilot: Population Based Cohort Study. *British Medical Journal*, 374:n1637.

Goldberg, S. (2010). *Relying on Others: An Essay in Epistemology*. Oxford University Press.

Goldman, A. (1976), Discrimination and Perceptual Knowledge. *Journal of Philosophy*, *73*, 771–91.

Goldman, A. (1979). What Is Justified Belief? In G. S. Pappas (Ed.), *Justification and Knowledge* (1–25). Reidel. Reprinted in A. I. Goldman, *Reliabilism and Contemporary Epistemology* (29–49), Oxford University Press, 2012.

Goldman, A. I. (1986). *Epistemology and Cognition*. Harvard University Press.

Goldman, A. and B. Beddor (2021). Reliabilist Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 ed.). https://plato.stanford.edu/archives/sum2021/entries/reliabilism/

Graham, P. (2014a). The Function of Perception. In A. Fairweather (Ed.), *Virtue Scientia: Bridges between Virtue Epistemology and Philosophy of Science* (13–31). Synthese Library.

Graham, P. (2014b). Functions, Warrant, History. In A. Fairweather and O. Flanagan (Eds.), *Naturalizing Epistemic Virtue* (15–35). Cambridge University Press.

Kitcher, P. (1992). The Naturalists Return. *The Philosophical Review*, *101*(1), 53–114.

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

Pernu, T. K. (2009). Is Knowledge a Natural Kind? *Philosophical Studies*, *142*(3), 371–86.
Pritchard, D. H. (2009), *Knowledge.* Palgrave Macmillan.
Sosa, E. (1992). Generic Reliabilism and Virtue Epistemology. *Philosophical Issues*, *2*, 79–92.
Stich, S. P. (1985). Could Man Be an Irrational Animal? *Synthese*, *64*(1), 115–35.
Tollefsen, D. P. (2009). Wikipedia and the Epistemology of Testimony. *Episteme*, *6*(1), 8–24.
Tharwat, A. (2021). Classification Assessment Methods. *Applied Computing and Informatics*, *17*(1), 168–92.

Appendix

- $\text{Precision} = \dfrac{\text{TP}}{\text{TP} + \text{FP}} \mid \text{High Precision} \Rightarrow \text{TP} \gg \text{FP}$

Manifesting high Precision is a distinct kind of reliability because it does not entail any of the following kinds of reliability:

- TN >> FN: That if the relevant mechanism classifies an element as a negative (e.g., not a cat), then it is highly likely that it is a negative indeed. For example, in the case of C, the mechanism might be prone to confusing (strange-looking) cats with objects that are not cats.
- TP >> FN: That the mechanism will identify most existing positives (e.g., cats). For example, in the case of C, there might be lots of cats in the vicinity, which C fails to classify as such (because, say, they are strange-looking cats). Consequently, a large percentage of positives will be ignored.
- TN >> FP: That the mechanism will identify most existing negatives (e.g., objects that are not cats). In the case of C, for example, even if the mechanism does not *often* misidentify objects that are not cats as cats (due to high Precision), the actual number of FP can still be close to the number of TN. This will be the case every time the number of positives in the population is very high and, thus, the number of negatives very low. In this type of case, the number of negatives that can be correctly identified (i.e., the number of TN) will be comparable to the number of FP—even if Precision is high.[26] Consequently, a large percentage of objects that are not cats will be ignored.

- $\text{Negative Predictive Value (NPV)} = \dfrac{\text{TN}}{\text{TN} + \text{FN}} \mid \text{High NPV} \Rightarrow \text{TN} \gg \text{FN}$

Manifesting high NPV is a distinct kind of reliability because it does not entail any of the following kinds of reliability:

---

26. When the number of positives is very high, high Precision (TP >> FP) does not mean that the actual number of FP will be low, especially when compared to N and TN.

- High Precision (TP >> FP): That if the mechanism classifies an element as a positive (e.g., a cat), then it is highly likely it is a positive indeed. For example, in the case of C, the mechanism might be prone to misidentifying as cats objects that are not cats (because, say, they are deceptively cat-like).
- TN >> FP: That the mechanism will identify most existing negatives (e.g., objects that are not cats). For example, in the vicinity of C, there might be lots of objects that are not cats but C fails to classify them as such (because, say, they look deceptively cat-like). Consequently, a large percentage of negatives will be ignored.
- TP >> FN: That the mechanism will identify most existing positives (e.g., cats). In the case of C, for example, even if the mechanism does not *often* misidentify cats as objects that are not cats (due to high NPV), the actual number of FN can still be close to the number of TP. This will be the case every time the number of negatives in the population is very high and, thus, the number of positives very low. In this type of case, the number of positives that can be correctly identified (i.e., the number of TP) will be comparable to the number of FN—even if NPV is high.[27] Consequently, a large percentage of cats will be ignored.

- $\text{Recall} = \dfrac{\text{TP}}{\text{P}} = \dfrac{\text{TP}}{\text{TP} + \text{FN}} \mid \text{High Recall} \Rightarrow \text{TP} \gg \text{FN}$

Manifesting high Recall is a distinct kind of reliability because it does not entail any of the following kinds of reliability:

- High NPV (TN >> FN): Even if the mechanism rarely misidentifies positives as negatives (due to high Recall it correctly identifies most positives), the number of objects that are correctly identified as negatives might not be much higher. This could be for two reasons. The mechanism may be likely to misidentify negatives (e.g., cat-like objects) as positives. Or, perhaps, the overall number of negatives (and thus the number of negatives that could be correctly identified) is low. In either case, the likelihood of a negatively identified element to be a negative indeed will be low.[28]
- High Precision (TP >> FP): That if the relevant mechanism classifies an element as a positive (e.g., a cat), then it is highly likely it is a positive indeed.

---

27. When the number of negatives is very high, high NPV (TN >> FN) does not mean that the actual number of FN will be low, especially when compared to P and TP.

28. NPV = TN / (TN + FN). Therefore, high NPV entails that TN >> FN. Due to high Recall, the mechanism will output very few FN. Nevertheless, the mechanism may still exhibit low NPV, because it outputs very few TN. Since N = TN + FP ⇒ TN = N - FP, this could be for two reasons. Either the number of FP is very high or the number of existing N is very low.

For example, in the case of C, the mechanism might be prone to misidentifying as cats objects that are not cats (because, say, they are deceptively cat-like).

- TN >> FP: That the mechanism will identify most existing negatives (e.g., objects that are not cats). For example, in the vicinity of C, there might be lots of objects that are not cats but C fails to classify them as such (because, say, they look deceptively cat-like). Consequently a large percentage of negatives will be ignored.

- $\text{Specificity} = \dfrac{\text{TN}}{\text{N}} = \dfrac{\text{TN}}{\text{TN} + \text{FP}} \mid \text{High Specificity} \Rightarrow \text{TN} \gg \text{FP}$

Manifesting high Specificity is a distinct kind of reliability because it does not entail any of the following kinds of reliability:

- High Precision (TP >> FP): Even if the mechanism rarely misidentifies negatives as positives (due to high Specificity it identifies most negatives), the number of objects that are correctly identified as positives might not be much higher. This could be for two reasons. The mechanism may be likely to misidentify positives (e.g., strange looking cats) as negatives. Or, perhaps, the overall number of positives (and thus the number of positives that could be correctly identified as such) is low. In either case, the likelihood of a positively identified element to be positive indeed will be low.[29]
- High NPV (TN >> FN): That if the relevant mechanism classifies an element as a negative (e.g., a non-cat), then it is highly likely it is a negative indeed. For example, in the case of C, the mechanism might be prone to misidentifying as objects that are not cats objects that are cats (because, say, they are strange-looking cats).
- High Recall (TP >> FN): That the mechanism will identify most existing positives (e.g., cats). For example, in the vicinity of C, there might be lots of objects that are cats but C fails to classify them as such (because, say, they are strange-looking cats). Consequently, a large percentage of positives will be ignored.

---

29. Precision = TP / (TP + FP). Therefore, high Precision entails TP >> FP. Due to high Specificity, the mechanism will output very few FP. Nevertheless, the mechanism may still exhibit low Precision, because it outputs very few TP. Since P = TP + FN ⇒ TP = P - FN, this could be for two reasons. Either the number of FN is very high or the number of existing P is very low.