

Duhemian good sense and agent reliabilism

Famously, according to Duhem a hypothesis can never be experimentally tested in isolation, but only along with the entire theoretical scaffolding it comes with. So in the face of disagreement between theory and experiment, it is impossible to point out which hypotheses in the theory are flawed. A big question for Duhem was, how does the physicist act in such a situation of underdetermination? Which hypotheses does s/he discard, and which one(s) does s/he retain? Duhem's response was that the physicist possesses an intuitive "good sense" that directs this choice. Although good sense does not provide a rigorous, rule-based template for theory choice¹, it allows scientists to weigh evidence and be "fair and impartial" (Duhem, 218) in theory choice.

Recently, there has been much interest in drawing parallels between Duhem's good sense and ideas in virtue epistemology (VE). VE emerged in the 1980s as an approach to epistemology based on virtue ethics. In the words of Greco (2004): "Just as virtue theories in ethics try to understand the normative properties of actions in terms of the normative properties of moral agents, virtue epistemology tries to understand the normative properties of beliefs in terms of the normative properties of cognitive agents." A virtue epistemological reading of good sense as first advanced by David Stump (2007) is based on the idea that Duhem too emphasized the normative properties of the scientist qua cognitive agent and took them as a basis for legitimate scientific

¹ While "theory choice" today is generally understood in the context of contrastive underdetermination, Duhem was primarily concerned with the holist variety of underdetermination and advanced good sense in the context of the latter. But for the purposes of this paper the distinction will not matter, and I shall use "theory choice" to refer to underdetermination in general, as do all the authors I reference.

knowledge in the face of underdetermination of theory by evidence. Stump finds striking similarities particularly between Duhemian good sense and Linda Zagzebski's (1996) views of VE. Here, I discuss the views of Stump, Milena Ivanova (2010), and Abrol Fairweather (2012) in this regard and ultimately propose my own view in response which is an agent-reliabilist reading of Duhem's good sense.

Stump argues that Duhem conceived of good sense in a way that can today be understood as virtue theoretic. In particular, Stump finds similarities between good sense and ideas of VE put forward by Zagzebski (1996). As Stump tells us, Zagzebski argued that justified belief comes from a "cluster of intellectual virtues in the same way that the rightness of an act can be defined in terms of moral virtue in ethical theory" (Stump, 151). Stump argues that Duhem's good sense nicely fits in with these ideas. Good sense depends on the scientist, the cognitive agent, being "virtuous": s/he has to be, in the words of Duhem quoting Claude Bernard, a "faithful and impartial judge". Stump further provides another illuminating quote from Duhem from his lectures on German science:

"In the realm of every science, but more particularly in the realm of history, the pursuit of the truth not only requires intellectual abilities, but also calls for moral qualities: rectitude, probity, detachment from all interest and all passions. (Duhem, 1991b, p. 43)" (Stump, p. 152).

Stump notes that some of the epistemic virtues put forward by Zagzebski include intellectual sobriety, impartiality and intellectual courage and the list fits very well with Duhem's. Yet another striking similarity between Zagzebski and Duhem according to Stump is that they both appeal to non-rule-governed epistemology. Zagzebski, in making a case for an

epistemology based on ethics, says, “The idea is that there can be no complete set of rules sufficient for giving a determinate answer to the question of what an agent should do in every situation of moral choice.” (Stump, 152) Similarly, Duhem arrives at the idea of good sense when the rule-based epistemology of the physical method (i.e. strict agreement between theory and experiment) fails. As Stump says,

“Holism threatens to make testing impossible, yet Duhem believes that scientific consensus will emerge. While the pure logic of the testing situation leaves theory choice open, good sense does not. Duhem claims that the history of science shows that while there is controversy in science, there is also closure of scientific debates.” (Stump, 155)

Milena Ivanova (2010) has argued in response to Stump, that the latter is mistaken in drawing such close parallels between VE and Duhem’s good sense. She raises two main objections: first, while VE is concerned with getting to the *truth* via epistemic virtues, for Duhem, physical theory only asymptotically approaches truth – truth here being the truth of a natural order, of the “real affinities” among things. Ivanova makes this point keeping in mind Duhem’s view of a ‘perfect theory’ and the convergent nature of his realism: for Duhem, the aim of physical theory was to classify experimental laws, and a physical theory – one picked out by good sense in the face of underdetermination – constantly approached but never reached, a perfect theory which classified laws and their phenomena in exactly the way underlying metaphysical realities are really classified in nature. So her point is that while VE is concerned with getting to the truth, good sense doesn’t help us with that. But as Ivanova herself points out,

“Still, in response to this objection one can adopt the weaker thesis that even though natural

classification may not reveal the truth about the unobservable, it will be true for the observable phenomena. Also, one may argue that it is legitimate to aim at a particular epistemic goal independently of whether this goal is achievable or not.” (62)

I take her point here to be that both VE and good sense are after all in the business of truth-seeking even though attaining the truth may be impossible for with the latter.

Ivanova’s more forceful objection has to do with epistemic justification. According to her whereas VE takes epistemic virtues to be *justifications* for beliefs, Duhem did not invoke the concept of good sense to *justify* belief in one theory over another. (To reiterate, Duhem did not have a full-blown metaphysical notion of truth of a theory – but worked with the surrogate idea of truth, that a right theory approaches a transcendental, natural classification.) Rather, she argues, good sense for Duhem was more a post hoc *explanation* of the physicist’s choice: it explains the repeated success of theories at making novel predictions. According to Ivanova, what really justified belief in a theory for Duhem – i.e. the belief that it was approaching a natural classification – was the success of the theory in making correct novel predictions: She says that for Duhem, “[a scientist] is justified in believing that a theory is a natural classification only when some empirical evidence supports it or when the theory has become a ‘prophet for us’ (Duhem, 27), that is, when it has managed to make novel predictions.” (Ivanova, 62). Here’s Ivanova’s argument broken down:

- Physical theory is a classification of laws.
- In a situation where we have a theory that contradicts experimental data and are left without any means within physics to decide what to do - whether to tweak parts of the theory to accommodate the available experimental data – and if so, which parts to tweak

- or to abandon it for another theory. Somehow in the end, the scientist decides which way to go.
- The “highest test” for physical theory is to ask it to make new and novel experimental predictions.
 - When the theory succeeds it is justified – in that it is taken to approach a natural classification.
 - Repeatedly, the scientist sees her/his choices made in the difficult situation of underdetermination emerging successful in such predictions.
 - How does this happen? There must be some innate ability or virtue in the scientist that enables him to do this: good sense.

Thus according to Ivanova, good sense is an explanation of theory choice rather than a justification for it. Moreover, according to her, Duhem doesn't say anything about good sense as a method of science: he doesn't tell us *how* exactly it directs our choice. His account of how good sense comes about and works to direct theory choice is quite thin. For Ivanova, this further shows that Duhem did not introduce it as a justification but only as a post hoc explanation.

Abrol Fairweather (2012) has argued against Ivanova's above objection and has attempted a position on Duhemian good sense that is a hybrid of Ivanova's and Stump's views. Fairweather claims to draw upon an agent reliabilist VE to do this. Reliabilism in Alvin Goldman's words, "... as a distinctive approach to knowledge is restricted to theories that involve truth-promoting factors above and beyond the truth of the target proposition." (Goldman, 2011) Fairweather's argument is that good sense results in a *reliable* process. Since Duhem's

claim is that good sense has a great “track record” and always picks out a successful theory – i.e. a theory which inevitably *correctly* makes a novel prediction – good sense produces knowledge (which here in the Duhemian context, consists in taking a predictively successful theory to be approaching a natural classification) by a *reliable* process. Good sense is a ‘truth-promoting factor’ regardless of whether the theory it picks out ultimately succeeds in novel prediction or not. It is “tracking evidentially important features of theories” (Fairweather, 10) Fairweather claims that “If a belief P is the product of a reliable capacity or process this fact constitutes evidence in favor of P.” This implies, “If the products of good sense reliably turn out to be supported by compelling new evidence, then being the product of good sense will be evidence for any theory with such a distinguished etiology.” (Fairweather, 10) So, Fairweather says, it seems that “future evidence is not required to evidentially distinguish the theory chosen by good sense, because the reliability of good sense is itself evidence supporting that theory.”

(Fairweather, 10) While I agree that agent reliabilism is the best way to understand good sense, Fairweather does not seem to give an accurate interpretation of this reading. Although he claims to provide an agent reliabilist reading of good sense, he grounds the reliability of good sense in its track record and not in its own nature or the mind where it is born. This is antithetical to agent reliabilist VE which situates reliability in the cognitive character of the agent. So it seems that Fairweather’s characterization is more along the lines of process reliabilism or simple reliabilism – according to which a belief is justified just in case it is formed via reliable processes – rather than agent reliabilism, and hence contrary to what he set out to do. His argument does not help situate good sense back into VE. Let us now turn to agent reliabilism in detail.

Greco and Agent Reliabilism: A Short Detour

As above, simple reliabilism is the view that a belief is justified just in case it is formed via reliable processes. Here the proportion of true beliefs the process results in, over time, measures reliability. Greco (1999) argues that simple reliabilism is insufficient for two reasons:

1. An agent might form a belief via fleeting or strange processes: Greco starts by noting that “Reliabilism must somehow restrict the kind of reliable process that is able to ground knowledge, so as to rule out processes that are strange or fleeting.” (Greco, 286) As an example of such processes, Greco discusses Platinga’s “The case of the epistemically serendipitous lesion” where an agent has a rare kind of a brain lesion, one that makes her believe that she has a brain lesion. There is no evidence for the lesion: there no symptoms, no testimony etc.; in fact there might even be a lot of evidence *against* it. But the agent is unable to take account of this (lack of) evidence due to the lesion. The relevant cognitive process here must no doubt be deemed very reliable, but we would not want to take the resulting belief as justified.
2. Process reliabilism doesn’t guarantee that the agent has a subjective justification of her belief. Greco says,

“[there] is a powerful intuition that knowledge does require that the knower have some kind of sensitivity to the reliability of her evidence. Sometimes this intuition is expressed by insisting that knowledge requires subjective justification. It is not enough that one's belief is formed in a way that is objectively reliable; one's belief must be formed in a way that is subjectively appropriate as well.” (285)

Greco’s solution to the above problems is agent reliabilism. According to agent reliabilism, reliability is shifted from the belief-forming process to the qualities of the agent’s

mind:

“Relevant to present purposes is Sosa's suggestion for a restriction on reliable cognitive processes; it is those processes that have their bases in the stable and successful dispositions of the believer that are relevant for knowledge and justification. Just as the moral rightness of an action can be understood in terms of the stable dispositions or character of the moral agent, the epistemic rightness of a belief can be understood in terms of the intellectual character of the cognizer.” (Greco, 287)

Following Sosa's views, Greco proposes that “knowledge and justified belief are grounded in stable and reliable cognitive character.”(Greco, 287) Accordingly, “We may now explicitly revise simple reliabilism as follows: A belief *p* has positive epistemic status for a person *S* just in case *S*'s believing *p* results from stable and reliable dispositions that make up *S*'s cognitive character.” (Greco, 287) Hence we see that reliability now has little to do with the truth of the resultant belief(s) but rather with the cognitive character of the agent.

Greco proceeds to show how agent reliabilism also solves the problem of subjective justification:

VJ: “A belief *p* is subjectively justified for a person *S* (in the sense relevant for having knowledge) if and only if *S*'s believing *p* is grounded in the cognitive dispositions that *S* manifests when *S* is thinking conscientiously.” (289)

By “thinking conscientiously”, Greco clarifies that he does not mean thinking with the purpose of finding truth, but rather the “usual state that people are in as a kind of a default mode – the state of trying to form beliefs accurately.” Greco contrasts this with epistemic “vices” such as trying to comfort oneself or trying to seek attention. Lastly, Greco points out that agent reliabilism reverses the “usual direction of analysis between virtuous character and justified

belief”. While non virtue theoretic epistemologies understand virtues in terms of justified belief, here justified belief is being cashed out in terms of virtues of the cognizer. “Virtuous belief is associated with the dispositions a person manifests when she is sincerely trying to believe what is true”, and “The dispositions that a person manifests when she is thinking conscientiously are stable properties of her character, and are therefore in an important sense hers.” (Greco, 290) Therefore, a belief formed this way will be subjectively appropriate.

Back to Duhem

Duhem’s views seem to exhibit all the features of agent reliabilism discussed above. In addition to the features of good sense and the physicist qua cognitive agent discussed so far I want to draw the reader’s attention to Duhem’s characterization of the different kinds of minds. For Duhem, the “strong and the narrow” mind is one capable of ordering and organizing laws and hypotheses into theories, and the “supple” mind or the “mind with finesse” – one capable of grasping a wide range of objects and at the same time able to group them logically – is the mind that produces good sense. This certainly seems to talk of “stable dispositions” in Greco’s sense of the term, that reflect the “cognitive character” of the scientist. Duhem takes pains to carefully describe the mind of the physicist and discuss beliefs and attitudes *in terms of* cognitive character traits and not the other way round. i.e. Duhem talks of legitimacy of beliefs in terms of cognitive character traits; he does not talk of the traits or “epistemic virtues” so to speak, in terms of the validity of beliefs. For instance, he says about those not interested in seeing a unified system of classification erected, “Only those who affect a hatred of intellectual strength were mistaken to the extent of taking the scaffolding for a completed building.” (Duhem, 103) There are several such instances where Duhem turns traditional non virtue-theoretic epistemology on its head and makes cognitive character traits basic. Now it remains to be seen if we can defend a view of

justification from good sense that goes with Greco's account. If we are successful in this, Ivanova's position will be untenable. Before going there though, let us return to Fairweather for a moment.

In addition to the argument from reliabilism, Fairweather advances another argument against Ivanova's "deflation of good sense": the position that good sense does not lend any epistemic strength or any justification to the chosen theory. The argument is that if good sense were indeed merely explanatory and post hoc as Ivanova claims, and not justificatory, then we are free to imagine a case where good sense doesn't intervene at all. After all, if good sense explains theory choice and there is no choice being made – i.e. no explanandum - we don't need an explanation. So let us suppose that we don't make any choice and just wait for a future novel prediction to make a choice and justify it. This might not be the most efficient way to choose a theory, but let us assume we do this nevertheless – for according to Fairweather, Ivanova's objection should imply the possibility of this solution. Fairweather rightly points out that in this situation we *might* again end up with an underdetermination: what if all competing theories pass the novel prediction test? Therefore, Fairweather argues, good sense must play an important epistemic role above mere explanation, in the face of such a "second level" underdetermination. But he goes further than that and says that without it, we *would* never end up with a determinate choice, even with new confirming evidence. What Fairweather is ignoring here is that future evidence *could* pick out a theory, however small the probability. It is possible that when all the options resulting from underdetermination are asked to make a novel prediction, only one succeeds, hence obviating the need for any further theory revision. But the important point is that good sense enters the scene even before such an attempt to single out a theory based on novel prediction. So the merit of good sense in my view does not lie in the inability of novel

predictions to single out a theory. It is more fundamental than that. But reasons for meriting good sense apart, let us again look at Fairweather's take on *what* the merit of good sense is.

According to Fairweather, good sense confers *uniqueness* to a theory (which, according to him, no future evidence can confer). But after good sense has uniquely picked out a theory, it is a successful novel prediction that counts as evidence in favor of the chosen theory. Fairweather makes the following interesting observation that follows from such a reading of good sense:

“This shows an interesting fact that new evidence in favor of a theory gives it a different epistemic standing depending on whether we are considering it alongside or independent of meaningful rivals. In the former case, new confirming evidence does not make a theory the determinate choice with fundamental epistemic standing. In the latter case, that same evidence determines theory choice and confers fundamental epistemic standing.” (Fairweather, 13)

So there are two “epistemic values and epistemic standings”: uniqueness, which comes from good sense, and clinching evidential support from a successful novel prediction. This way, good sense alone does not confer “fundamental epistemic standing”, and evidence alone cannot confer uniqueness. This account which recognizes an important epistemic role for both good sense and new evidence, Fairweather calls the “hybrid reading”.

My own view is that while Fairweather is right in that good sense plays a key epistemic role unlike what Ivanova says, we can go back full circle to Stump and have a proper virtue epistemological – specifically agent reliabilist – reading of good sense. I contend that good sense confers not just uniqueness, but actually does determine theory choice, also providing (an agent-

reliabilist) justification. Good sense doesn't simply pick one and put the rest "out of the running". It is not just something that prevents the proliferation of acceptable theories obtained by tweaking different parts of theories that don't agree with future experiment. Good sense provides a *basis* for the uniqueness. Just as with the problem of coming up with a realist interpretation of Duhem, this problem of the epistemic role of good sense is not easy either given the sometimes confusing nature of Duhem's claims. Nonetheless, I still think an agent-reliabilist VE reading of Duhem is possible and that Ivanova and Fairweather are mistaken.

Ivanova claims that good sense is only offered as a post hoc explanation of theory choice during underdetermination and not as a justification. I argue to the contrary. Ivanova's claim seems to be based on a purely externalist notion of justification. It seems to assume that there is one single concept of justification – specifically, externalist, evidential – and that good sense doesn't fit with it. But justification can be of many kinds. Duhem says we can "very properly decide" (Duhem, 217) between multiple theory choices using good sense. Further, he says good sense strongly "comes out in favor of" one of the choices – again implying that we are compelled to accept its judgment *even before* future experiment can ratify the choice. He goes on to say, "Pure logic is not the only rule for our judgments; certain opinions which do not fall under the hammer of contradiction are in any case perfectly unreasonable." (Duhem, 217) How do we understand such language? If an epistemic choice is proper, forceful, and reasonable, I don't see any reason we cannot properly construe it as being justified, in an *internalist* sense.

Further, Duhem does *not* introduce good sense as a merely post hoc explanation. He says, we can "properly decide" between the various options of theories using good sense. "Properly

decide” very much implies an active role for good sense *during* underdetermination. Duhem presents elaborate and careful characterizations of different kinds of minds and puts forward quite clearly, *normative* merits of cultivating/ possessing one kind of mind over the other as far as physics goes (the supple or the strong and narrow over the ample, broad and weak). Good sense is but a feature of the supple mind. It is not introduced all of a sudden as a new idea to just “save the (meta)phenomenon” of theory choice during underdetermination. It is a smooth and natural continuation of Duhem’s views on the mind of the theorist, which he articulates way before he comes to this problem of underdetermination, in one of the early chapters in *Aim and Structure*. In fact, Duhem’s view that physicists don’t actually actively choose hypotheses at all, and that they “come to his mind” when his mind is ready to receive them, clearly reveals the agent reliabilist in Duhem.

Finally, Greco’s account of agent reliabilist justification seems to lend itself to Duhem very well. Reliable cognitive character *justifies* beliefs it produces and further, it is subjectively justified: Duhem’s virtuous scientist certainly “thinks conscientiously”, following Duhem’s instructions of shunning passions and interests, and so a belief, here the belief in the theory chosen, grounded in the cognitive dispositions, here good sense, he manifests when thinking like this – is subjectively justified. So we seem to have comfortably accommodated Duhem in a full-blown agent reliabilist reading.

But what about the textual evidence cited by Ivanova, which seems to say Duhem did not think good sense justified theory choice? Why does Duhem insist that despite good sense, it is a successful novel prediction that has the final word? Why does he, in the context of resolving

underdetermination say in as many words that the method of the physicist “is justified only by experiment”? I contend that throughout *Aim and Structure*, Duhem seems to have two distinct, non-intersecting epistemologies: one of physics, and one outside of physics – which we may call philosophy. Duhem was a physicist-philosopher. He frequently claims that although there are absolutely no epistemic resources *within* physics for us to believe that physical theory latches on to a natural underlying order, we are forced to believe so by various factors outside of physics, logic and reason. It is worth noting that Duhem cites Pascal as saying that we sometimes believe for ‘reasons that reason does not know’, both in the context of theories converging on to a natural classification as well as in that of good sense during underdetermination. About the former, he says: “The opinion is a legitimate one because it results from an innate feeling of ours which we cannot justify by purely logical considerations, but which we cannot stifle completely either.”

(Duhem, 102) Further:

“No language is precise enough and flexible enough to define and formulate them; and yet, the truths which this common sense reveals are so clear and so certain that we cannot either mistake them or cast doubt on them; furthermore, all scientific clarity and certainty are a reflection of the clarity and an extension of the certainty of these common-sense truths.” (Duhem, 104)

Since Duhem attributes good sense to similar patterns of thinking, we can associate his above assertions about the legitimacy of beliefs not borne out of logic, with good sense as well. Given Duhem’s commitment to the moral goodness and the intellectual acuity of the supple, strong and narrow minds, it is very unlikely that he would think that epistemic ends justify the means (here, successful novel prediction justifying that which chose the theory, i.e. good sense). Reliabilism in fact expressly turns this around and say it is the means (by virtue of their

reliability) that justify the ends. So beliefs that arise from good sense are *justified* from an (internalist, deontological) agent reliabilist perspective. The justification Duhem talks about when he says that the methods of the physicist are justified by experiment should be when we are strictly within the context of physics: there it is Duhem qua physicist speaking. But from a broader, philosophical perspective, Duhem rather means, I think, that experiment *validates* the choice and confers *certainty* on it. But we can have justification without certainty, like in agent reliabilism. In simpler terms, the *reasons* for which the physicist chooses a theory are grounded in her good sense. However, the successful novel prediction will no doubt make the choice certain.

Thus, Ivanova is mistaken in arguing that good sense does not provide justification. Fairweather's hybrid reading is inadequate as well for it ignores the justification offered by a proper agent reliabilist reading of good sense. I argue that a proper agent reliabilism accommodates Duhem as a virtue epistemologist very well and shows us that good sense does offer justification for theory choice. Importantly, I have shown that it is certainly not a post hoc explanation but a part and parcel of Duhem's overall views on the mind of the physicist.

References

- Duhem, Pierre. (1954). *The aim and structure of physical theory*. Princeton: Princeton University Press.
- Fairweather, A. (2012) 'The Epistemic Value of Good Sense' *Studies in the History and Philosophy of Science* <http://philpapers.org/archive/FAITEV.pdf>

- Goldman, Alvin. 'Reliabilism', *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/reliabilism/>.
- Greco, J. 1999. 'Agent reliabilism' in *Philosophical Perspectives* 13: 273-296.
- Ivanova, M. (2010). 'Pierre Duhem's good sense as a guide to theory choice'. *Studies in History and Philosophy of Science*, 41, 58–64.
- Stump, David. (2007). Pierre Duhem's virtue epistemology. *Studies in History and Philosophy of Science*, 38, 149–159.