Non-citable postprint. Forthcoming in *Asian Journal of Philosophy*, July 2024.

Epistemic Consequentialism as a Metatheory of Inquiry

Frederik J. Andersen & Klemens Kappel

Forthcoming in *Asian Journal of Philosophy*

Abstract

The overall aim of this article is to reorient the contemporary debate about epistemic consequentialism. Thus far the debate has to a large extent focused on whether standard theories of epistemic justification are consequentialist in nature and therefore vulnerable to certain trade-off cases where accepting a false or unjustified belief leads to good epistemic outcomes. We claim that these trade-offs raise an important—yet somewhat neglected—issue about the epistemic demands on inquiry. We first distinguish between two different kinds of epistemic evaluation, viz., *backing* evaluation and *outcome* evaluation, and then go on to outline and discuss a consequentialist metatheory about the right combinations of decision procedures to adopt in inquiry. Note that the piece is exploratory in the following sense: we try to explore epistemic evaluation in consequentialist terms, which involves stating a form of epistemic consequentialism, but also pointing to what non-consequentialist alternatives might be. Rather than trying to argue decisively for a particular conclusion, we aim to outline various intricate issues in an underexplored area of theorizing. In the course of doing this, we'll transpose some well-known themes from discussions of consequentialism in ethics to the current debate about consequentialism in epistemology, e.g., agent-neutrality, options, and side-constraints.

Keywords

Epistemic Consequentialism; Epistemic Backing Evaluation; Epistemic Outcome Evaluation; Inquiry; Criteria of Rightness; Decision Procedures; Trade-offs

1 Introduction

A current topic in epistemology is whether our standard theories of epistemic justification are consequentialist, i.e., do they bear a strong resemblance to the well-known consequentialist theories in ethics. On one side of this debate Selim Berker has argued that a "consequentialist specter" haunts both ethics and epistemology, meaning that orthodox theories from both fields are consequentialist and thus vulnerable to the same types of objections including unacceptable trade-offs.[1] In several papers (2013a, 2013b, 2015) he has criticized a wide range of theories—among these *process reliabilism*—for being consequentialist and for that very reason implausible.[2] On the other side, various authors have defended some of these supposedly consequentialist theories, e.g., Kristoffer Ahlstrom-Vij and Jeffrey Dunn (2014, 2017) who concede that process reliabilism is a consequentialist theory in a certain sense, though not one that is vulnerable to the objections that Berker raises.[3]

Let's start out by considering the kind of trade-off cases Berker and others have found problematic for process reliabilism:

> ***The Truth-Fairy***. Suppose a truth-fairy exists who guarantees that, if I believe that Lake Constance is filled with milk—a proposition I have many good reasons to think is false and that is, in fact, false—then she will arrange things in a way that all the other (and logically independent) beliefs I entertain now and in the future will be true. Given these circumstances, should I believe that Lake Constance is filled with milk? Would that belief be epistemically justified? (Briesen 2016, p. 288)[4]

A host of similar cases can be found in the literature, all featuring more or less fanciful imaginaries making it true that if an agent adopts a belief for which there is no good evidence (or other epistemic backing), then overwhelmingly good epistemic consequences will be the result. Critics of consequentialism claim that in such cases process reliabilism sacrifices one proposition for the sake of others, and that this is epistemically unacceptable. See for instance (Berker 2015).

[1] Berker isn't alone in criticizing epistemic consequentialism. Others have criticized epistemic versions of consequentialism (and/or similar views). See for example (Carr 2017; Friedman 2019; Greaves 2013; Kelly 2003; LittleJohn 2012, 2018). Note also the relevance of Friedman's (2020) seminal discussion of *zetetic* epistemology here.
[2] For canonical formulations of process reliabilism, see for example (Goldman 1979, 1986).
[3] See also (Andersen and Kappel 2020; Goldman 2015).
[4] See also (Fumerton 1995; Jenkins 2007).

For the purposes of this paper we want to set aside the issue of whether trade-off cases of the mentioned kind pose a genuine problem for process reliabilism (or any other theory of epistemic justification for that matter). Instead we want to suggest that these cases raise an important but somewhat neglected issue. We submit that a distinctive feature of the relevant kind of trade-offs is that they prompt both epistemic *backing* evaluation and epistemic *outcome* evaluation. The former kind of evaluation concerns the epistemic backing of a target belief at a given time, where "backing" can be thought of in several different ways: evidentialism, process reliabilism, virtue responsibilism etc. The latter kind focuses on epistemic outcomes instead, i.e., on what results from an agent's adopting a given target belief (epistemically speaking). Note that this distinction shouldn't be thought of as being temporal in any crucial way. For example, it's not the case that outcome evaluations are essentially future-directed (or forward-looking), while backing evaluations are bound by the moment (or sideward-looking). Of course, we can only evaluate a given belief's epistemic backing (evidence, reliable and/or virtuous formation etc.) relative to a certain point in time. But such epistemic backing at such time could depend upon, or be constituted by, facts or processes that are temporally prior to it. When it comes to epistemic outcome evaluation, it's also clear that this is basically a matter of evaluating a target belief by assessing what epistemically relevant "stuff" results from an agent's adopting the belief in question. This stuff can take the form of later beliefs, and typically it will, as in the **Truth-Fairy**. But clearly it need not be so. The crucial matter is the relationship between a target belief—which is being evaluated—and the epistemically relevant items that enter the evaluation *as outcomes* of adopting the belief.

When reflecting a bit on these remarks, it becomes obvious that outcome evaluation can (and should) be extended to other items than beliefs. Actions, processes, dispositions, methods, and patterns of inquiry, are all fitting candidates for being epistemically outcome evaluated. Stated more rigorously we want to suggest that:

An *epistemic backing evaluation* assesses whether a doxastic attitude (such as outright belief that *p* or suspension of judgement as to whether *p*) is sufficiently supported by S's total evidence—or by some other epistemic backing like the reliability of the cognitive process that gave rise to the attitude in question—at time t.[5] [6]

[5] For a canonical formulation of evidentialism, see (Feldman and Conee 1985).

[6] Note that we use the symbol 'S' to denote an arbitrary subject (alternatively: agent, reasoner, cognizer etc.) throughout the paper. Moreover, we use symbols 't' and '*p*' to refer to a specific point in time and a specific proposition, respectively. Note, finally, that we use the expression 'epistemic backing' as an umbrella term meant to subsume the focal points of various standard theories of epistemic evaluation, i.e., evidentialism, process reliabilism, virtue responsibilism etc.

An *epistemic outcome evaluation*, by contrast, assesses a doxastic attitude (such as belief that *p*) or some other possibly non-mental evaluable item (like the act of gathering evidence) not solely based on its standing vis-à-vis S's total evidence (or epistemic backing) at t, but also its relation to strategic outcome-considerations which *can* involve future or prior states.

For example, an unsupported (false) belief might be epistemically right from a perspective of outcome evaluation if holding the belief enables you to reach a certain goal of inquiry in the long run. Epistemic justification—as it is normally thought of in mainstream epistemology—merely concerns backing evaluation, but even so, outcome evaluations do make intuitive sense, and in trade-off cases it's clear that backing and outcome evaluations can come apart. In the **Truth-Fairy**, the belief that <Lake Constance is filled with milk> is not epistemically justified (qua backing), and yet, from an outcome perspective there is something epistemically right (or at the very least permissible) about it.

Here, we want assert that it's indeed natural to evaluate *inquiries* from an epistemic outcome perspective, i.e., a perspective from which the epistemic assessment of some item, open for evaluation, isn't solely based on its current epistemic backing, but also its relation to wider strategic outcome-considerations such as garnering important true beliefs in a near future, accepting provisional (perhaps implausible) idealizations in mathematical models in order to gain a better grasp of an empirical phenomenon, or playing a certain role in a team of inquiring agents with a common epistemic goal.[7] The main purpose of this paper is to illustrate that when it comes to the epistemic demands on inquiry, i.e., what one ought to do epistemically speaking given one's aim of inquiry, it's plausible that we should be consequentialists about which combinations of decision procedures to adopt. For the rest of the paper we'll assume that *inquiry* is an activity organized around the aim of reaching a certain epistemic goal, namely maximizing epistemic value with respect to some target proposition at the end of inquiry. In other words, we'll take inquiries to be proposition-relative. As the paper is explorative, we'll not take a firm stance with respect to epistemic value theory, but simply note that true belief, justified belief, rational belief, knowledge etc., are all live candidates of (basic) epistemic goods.

Finally, before ending the introduction, let's have a quick look at some illustrative cases of inquiry that will foreshadow a number of important themes for us, e.g., the tight connection between

[7] Note that Brian Talbot (2014) has presented a distinction between *evidential* and *truth promoting non-evidential* reasons for belief, which is similar, though not identical, to our division between epistemic backing evaluation and epistemic outcome evaluation. See also (Conee 1987) for a similar distinction.

inquiries and epistemic outcome evaluation; and the stark contrast to mere epistemic backing evaluation.

> ***Surveying Literature***. In my inquiry about a certain problem—i.e., the truth value of target proposition *p*—I need to read through a large amount of literature of varying quality. As it happens, there is no other way I can achieve my epistemic goal, which is knowing the truth value of *p*. Although this procedure will eventually get me to the right answer to my question, I will inevitably adopt a number of false beliefs along the way. So, these false beliefs are a subset of the set of beliefs that comes with the best path of inquiry-steps available to me at this moment. My doing such a literature survey seems epistemically right when assessed in terms of outcome.
>
> ***Unjustifiedly Believing a Theory***. To gain knowledge about some important target proposition, *p*, I need to understand a theory T which is both complicated, counterintuitive, and controversial. In order to understand the theory, I will need to convince myself that it is true, otherwise I cannot really make myself grapple with it and fully grasp it. To do so, I see a teacher whose strong and convincing personality causes me to believe the theory, though neither the testimonial evidence from the teacher nor my other epistemic backing makes me epistemically justified in doing so. Thus, I am unjustified in my belief that <T is true>, yet this belief is a psychological precondition for gaining knowledge about my target proposition of inquiry, *p*, and thus it is a necessary inquiry-step for me to take in order to reach my epistemic goal. Hence, my unjustified belief that <T is true>, seems epistemically right when assessed in terms of outcome.
>
> ***Joint Inquiry***. A vicious crime happened Tuesday night. Art and Beau are detectives jointly trying to solve this criminal case by coming to know whodunit. Jones is the most likely suspect, so the most promising inquiry-step to take at this moment would be to try to determine what Jones was doing Tuesday night (e.g., via face-to-face interrogation of the prime suspect). However, Smith is also a suspect, though a slightly less likely one. Now, if one of the detectives had been on their own and only been able to investigate one of the two suspects' whereabouts, it would have been *prima facie*

epistemically irrational for that particular detective to investigate Smith (the less likely suspect), but as a team Art and Beau decide to divide their attention, so that Art looks at Jones' whereabouts Tuesday evening, whereas Beau investigates Smith. Assuming that Jones and Smith are almost equally likely to have committed the crime and that the investigation of either one of the suspects is a manageable job for just one detective in the given circumstances, it seems perfectly rational for Art and Beau to make this move as a part of their overall strategy. That is, it seems that it can be epistemically right (from a perspective of outcome evaluation) for the detective team to divide their attention and let Beau spend his resources on a path of inquiry, which considered on its own is less likely to lead to the truth.[8] Indeed, such a division of cognitive labour seems right when it is a part of the best currently available sequence of inquiry-steps.

Evidence Selection. You are a doctor who needs to make up your mind about whether some medical intervention actually works to improve the survival of your patients. You aim for a true belief, but unfortunately there is far more scientific papers about the intervention than you can survey. Some are arguing for the efficacy of the intervention, while others argue against it. Moreover, the published papers are of varying quality, ranging from very poor to very high. You are unable to provide a detailed assessment of the quality of more than a few of those papers, as this would be too complicated and take too much time. As it happens, a small subset of the papers report randomized controlled studies. Moreover, it is easy to identify those papers: a quick glance at the abstract will tell you whether a paper reports a randomized controlled study or not. Initially you decide to look only at the randomized controlled studies. It's not that by looking solely at those studies you will avoid all studies of poor quality; some will be good and some not. Some of the studies that you ignore might even report the truth about the efficacy of the medical intervention, and offer good evidence for it. Yet, your deliberate neglect of perfectly good evidence seems to be epistemically right (from a perspective of outcome evaluation) as long as this is part of the best available inquiry-path at the time.

[8] This is adapted from Kitcher's discussion of division of cognitive labour in science (Kitcher 1990).

The distinctive feature of these cases that we want to emphasise is that in order to achieve certain epistemic goals agents must countenance certain epistemic costs, and this may involve adopting false or unjustified beliefs, pursuing evidence that is not the most promising, or knowingly ignore evidence of high quality. There are obvious reasons why such trade-offs are common in everyday inquiry as well as scientific inquiry. Most of our non-trivial beliefs are the result of complex processes of inquiry. Inquiries often involve surveying large amounts of intricate and conflicting evidence and adopting auxiliary beliefs that enable further inquiry; sometimes they require careful distribution of cognitive efforts as well as strategic decisions about what evidence to seek out and what to ignore. Since these predicaments are common, the trade-off situations illustrated are as well.

2 Epistemic Consequentialism and Decision Procedures of Inquiry

To be sure, consequentialism is standardly understood as a *criterion of rightness* rather than a *decision procedure*. In ethics, a criterion of rightness specifies the necessary and sufficient conditions for an action to be morally right (or permissible); whereas a decision procedure is some trait, disposition, method, rule, heuristic etc. (or combinations of these), that agents use more or less successfully for determining what action(s) they ought to perform in a given situation (Parfit 1984). Similarly, in epistemology, one might think of epistemic consequentialism as a criterion of rightness specifying what it is for a belief (or other doxastic attitude) to be epistemically right, whereas a decision procedure is some habit, character trait, belief-forming disposition, method, rule or precept, heuristic, process, etc. (or combinations of these) that agents use (more or less successfully) for determining what belief(s) (or other doxastic attitude(s)) to form in a given situation.

As will become clear in the following, we want to highlight a form of epistemic consequentialism that is neither a decision procedure for doxastic attitudes, nor a first-order criterion of rightness for doxastic attitudes. Instead, we want to focus on epistemic consequentialism as a *higher-order criterion* that is used to evaluate combinations of decision procedures. Again, in parallel to ethical theory, one can individuate decision procedures in a very broad sense. In the context of inquiry this includes belief-forming dispositions telling us whether or not to adopt new beliefs, rules for revising existing doxastic attitudes like credences, precepts to look for new evidence or re-assess old evidence; and heuristics about what evidence to gather, and what methods to employ when doing

it.[9] Understood in this way decision procedures can obviously be of very different sorts. Some decision procedures are clear rules or precepts that agents can consciously follow, while others are habits, character traits, or belief-forming dispositions of individuals. Oftentimes, important decision procedures are integrated into social practices as we see it in science, various legal systems, media etc. Decision procedures may also vary vastly in levels of generality. Some apply to a broad range of decision problems, others to much more specific problems. Further, some decision procedures are transparent, say, when they consist of rules that one can consciously decide to follow. Other decision procedures are less transparent, and may consist of various practices that we follow without much reflection. Some decision procedures can be applied at will whereas others are more hardwired in our brains, or built into social practices that are hard to change.

Now, epistemic consequentialism is certainly not plausible when viewed as a decision procedure. As is familiar from ethics, an ethical theory may well direct agents to avoid using that theory as a decision procedure in their everyday moral inquiries. The most notorious example of this is of course consequentialism, which in many of its forms recommends that agents don't normally employ consequentialism as a decision procedure (see further discussions in (Parfit 1984) and (Hare 1981)). Yet, various forms of consequentialism in ethics can be highly intuitive (and plausible), when viewed as mere criteria of rightness.

Though things are surely different in many ways in epistemology, we nonetheless want to claim that something similar is true of epistemic consequentialism. While epistemic consequentialism might be applied to many items, such as individual beliefs or steps in inquiry, we submit that epistemic consequentialism is an illuminating and plausible higher-order criterion telling us which decision procedures, or rather *combination* of decision procedures, we ought to use in inquiry. Epistemic consequentialism—when thought of in this way—is simply telling an agent to adopt the

[9] This very broad understanding of decision procedures will inevitably lead to issues of unclarity—albeit none of them worse than what is simply taken for granted in the analogous discussion in ethics, see for example (Hare 1981; Parfit 1984). In ethical theory authors typically express decision procedures in a vague fashion, leaving it indeterminate what they imply in some cases. This would of course be a serious issue if we were specifying criteria of rightness, but it's not much of a concern when it comes to decision procedures. What ultimately matters is the outcome of using such procedures in practical decision-making, not whether they can be stated precisely. So, while it is certainly tempting to spell out the nature of decision procedures in much more detail, this would likely render them prohibitively difficult to use in practice. Decision procedures might work well in terms of the practical consequences of agents' complying with them even if they are very vaguely stated. Note also that Kantians and virtue ethicists may assert that part of what makes an action right is that the agent decides to perform the action in a certain way, or performs it with a particular intention, or instantiates certain character traits when performing it. On such views, what makes an action morally right is partly constituted by its being performed on the basis of a specific decision procedure. Kantians and virtue theorists should prefer decision procedures that constitutively make agents perform right actions. However, consequentialist and non-consequentialist may recommend the same decision procedures regardless.

best combination of decision procedures available to the agent,[10] where a combination of decision procedures is best if and only if no other combination of available decision procedures promote epistemic value better.[11] If this sounds almost like a truism, it may well be because epistemic consequentialism about decision procedures provides a genuinely plausible higher-order criterion of rightness applying to combinations of decision procedures. Of course, epistemic consequentialism doesn't directly provide us an operational *higher-order decision procedure* by which we can select the right combinations of decision procedures, and thus selecting the right ones can be a difficult task.

One might wonder why epistemic consequentialism should not be applied to individual beliefs or steps of inquiry, rather than decision procedures. The reason is that individual inquiry-steps, e.g., a revision of belief, do not have determinate consequences vis-à-vis the goal of inquiry independently of the sequence of steps they are part of. Compare to the evaluation of moves in a game of chess. Clearly, moves in a game of chess should be evaluated relative to the goal of winning, but a particular move can be evaluated only as part of a sequence of steps (or a set of possible sequences), also crucially depending on the "feedback" from the opponent. For the sake of illustration, the reader should recall the case of **Joint Inquiry** (cf. Section 1), where one of the two detectives takes an inquiry-step that would have seemed irrational from a purely individualistic backing perspective, and yet their division of cognitive labor seems epistemically right as a decision procedure given the social setting and mutual goal of inquiry. The chess-analogy reveals a fairly general feature of inquiries: any step in an inquiry can be epistemically evaluated only as part of a sequence of steps, which can be heavily influenced by one's (social) environment.

Thus, it makes sense for epistemic consequentialism to focus on the consequences of adopting decision procedures, or combinations of such, rather than the consequences of individual inquiry-steps. Note here that good decision procedures will be suited to the specific inquiries we face as well as the specifics of our cognition and patterns of social interaction that our inquiries may involve. Good decision procedures will counteract or productively use various cognitive biases we are subject to. Good decision procedures are, in some sense, possible to follow, adhere to, or adopt. This means that good decision procedures are consistent with the extent to which we can control our beliefs.

[10] Note that there are complicated issues about what it means for a decision procedure to be *available*. These issues are well-known from the debate about consequentialism in ethical theory, but to keep things simple we will just hint at the work of Erik Carlson here (Carlson 1995).

[11] Clearly, we take epistemic consequentialism to be a strict maximizing theory, see, e.g., (Brown 2011). Note, however, that some authors have suggested that consequentialists can allow for the permission of suboptimal alternatives, see for instance (Howard-Snyder and Norcross 1993; Slote 1984). Further, consequentialist theories come in *subjective* and *objective* versions. The former ranks alternatives in virtue of their *expected* outcomes, while the latter focuses on the *actual* outcomes. In this paper we'll assume epistemic consequentialism to be of the objective kind.

Doxastic voluntarism[12] is generally false, but we can indirectly affect what beliefs we adopt by controlling how we influence our cognitive system, e.g., what evidence we expose ourselves to. Good decision procedures should not make impossible demands. For example, as we usually don't have the information required, we cannot in general be expected to (directly) employ a rule like: *Adopt alternative X if and only if no other available alternative at this time has better epistemic consequences with respect to target proposition p*. In part this is why epistemic consequentialism is not plausible as a decision procedure. But, as we said, epistemic consequentialism merely tells us to adopt those combinations of available decision procedures which yield the best epistemic outcome given our aim of inquiry, it's silent as to *how* to do so in the heat of the moment.[13]

To provide a concrete and quite substantial example of the interaction we envision between different decision procedures and the proposed form of epistemic consequentialism, consider a quintessential question for epistemic virtue theory—viz., what makes some intellectual character traits epistemic virtues? Epistemic consequentialism provides a straightforward answer: *epistemic virtues are intellectual character traits that (when taken in isolation or in certain combinations) bring about the best epistemic consequences in inquiry*. When inquiry is taken to be proposition-relative, a particular intellectual character trait can of course be conducive to valuable epistemic ends in some inquiries and not in others, but there is nothing worrying or surprising about this if one goes for the consequentialist framework we have suggested. Epistemic virtues are just decision procedures, and the consequences of exercising them may vary depending on the details of one's informational and/or social setting as well as the possible combinations with other available procedures. We take this to be a significant example because epistemic consequentialism about decision procedures offers a plausible explanation of what makes some character traits virtues, and also a clear rationale for why we should care. Naturally there might be some very general traits that are thought to be epistemically good in all actual contexts, or even all possible contexts. Yet such general virtues may not be as significant as a host of responsibilists have suggested. Within the Aristotelian tradition the core virtues and vices have for the most part remained unchallenged. Interestingly, however, recent empirical findings in psychology may challenge this orthodoxy. Many psychological dispositions that are normally considered to be entirely virtuous have been found to correlate with increased polarization. For example, *open-mindedness*, normally taken to be a central epistemic virtue, has been

[12]*Doxastic voluntarism* claims that human agents can themselves—at least to some degree and with respect to certain types of beliefs—control what they want to believe. See for example (Alston 1988; Feldman 2001).
[13] Note that this is similar to, but not identical with, Goldman's suggested way of evaluating doxastic practices in his book Knowledge in a Social World (1999). One key difference is our explicit focus on *combinations* of decision procedures.

found to predict more extreme beliefs about climate change on both sides of the American debate, i.e., more open-minded conservatives are more likely to deny that climate change is occurring (Kahan and Corbin 2016). Clearly, a virtue such as open-mindedness will be epistemically right in some settings and not in others, if by epistemically right we mean conducive to sufficiently good epistemic outcomes in inquiry. The same holds for a vice such as being dogmatic. While it might often be epistemically wrong, it's not always so (Fantl 2021; Hallsson and Kappel 2018).

2.1 Non-Consequentialist Theories of Epistemic Outcome Evaluation

It's important to note that even our best combinations of available decision procedures are likely to be imperfect in at least one way, viz., they might have us adopt beliefs that are not epistemically justified qua our backing (or reject beliefs that are epistemically justified in this sense). This is interesting as it suggests a possible way in which a theory of epistemic outcome evaluation could be deontological or *non-consequentialist*: it could forbid adopting beliefs on insufficient backing, e.g., evidence, even when doing so would promote one's epistemic target of inquiry at least as good as any other available alternative at the time. For a belief to be epistemically right in the outcome-sense, it should at least be epistemically justified qua backing, a non-consequentialist might claim.

One could also imagine a non-consequentialist theory of epistemic outcome evaluation that permits *options*. In the present context, an option should be understood as a permission not to choose the best available alternative. So, a theory of outcome evaluation permitting options would have some combinations of decision procedures being permitted, although they do not promote the best outcome. This points to the fact that in devising a theory of epistemic outcome evaluation we face a question of whether one should accept unrestricted consequentialism, constrained maximization or maximization with some non-maximizing options permitted, and if the latter, what characterizes the relevant constraints and/or options.[14]

A related issue—also widely discussed in ethics—concerns *agent-relativity* versus *agent-neutrality* (Nagel 1989; Parfit 1984). Consequentialism in ethics standardly assumes that the values to be promoted are agent-neutral in some sense. There is, of course, a difference between whether an action benefits me or benefits someone else, but from a standard consequentialist point of view it counts just the same. The analogous assumption for epistemic consequentialism would be agent-neutrality with respect to epistemic value. At first sight, this may seem quite implausible. When I'm

[14] This is another question which has a familiar counterpart in ethical theory (Kagan 1998).

engaged in inquiry concerning target proposition *p*, rationality would seem to require that I do what best promotes a maximization of *my* epistemic value relative to *p*. So, I should maximize epistemic value in a way that is both agent-relative and proposition-relative. It seems that my identity and the focus of my inquiry determines what epistemic value to be pursued. Hence, it seems that a plausible form of epistemic consequentialism is wedded to a distinctive kind of agent-relativity and proposition-relativity regarding epistemic value.[15] In ethics, this very strong form of agent-relativity would amount to a form of ethical egoism. Most moral philosophers don't accept ethical egoism as it's rather difficult to make sense of the idea that it matters all that much for what I morally ought to do whether some benefit falls on me or on someone else, if everything else is equal. In the same vein one might wonder whether epistemic egoism is the most plausible form of epistemic consequentialism, after all. Sometimes a group of agents collaborate in inquiry concerning target *p*. Here, the optimal distribution of effort can be one in which not all agents individually aim to maximize epistemic value relative to *p*, say, because the most effective distribution of cognitive efforts requires agents to explore different avenues or test different sub-questions.[16] Even when inquiry is not collective in some organized sense, our inquiries often affect one another in similar ways. Why couldn't it be epistemically right in just the same sense to promote true beliefs or other epistemic ends in other subjects? If we concede that having true beliefs about questions that matter to us is epistemically valuable, why should it only be right to promote this value in our own lives and not in the lives of others? (For a recent discussion of related issues, see (Scott 2023)).

Note that this is not an issue that only affects hardcore consequentialist theories of epistemic outcome evaluation. Fully fledged theories accepting constraints and/or options (as discussed above) also need to take a stance vis-à-vis the agent-neutrality of epistemic ends. We'll not try to solve this intricate issue here as our aim is exploratory.

2.2 Trade-offs and Truth Demons

Finally, there is another important issue which we have neglected until now. As our framework is stated above, epistemic consequentialism is neither sensitive to the time nor other resources spent on an inquiry. If two combinations of available decision procedures both lead to the same degree of success with respect to the goal of inquiry, they are equally good, and therefore equally right, even if they are not equally expensive in terms of time or other resources invested in them. This is clearly

[15] Note that Goldman, in his framework for social epistemology, defines veritistic value in a way that makes it both agent-relative and proposition-relative (Goldman 1999, p. 89).
[16] Again, this is adapted from Kitcher's discussion of division of cognitive labour in science (Kitcher 1990).

not satisfactory: there is a sense in which it's suboptimal to choose a combination of decision procedures that's more demanding in terms of time and resources, if the result is the same.

It's not obvious how resources spent on inquiry should figure in a plausible consequentialist theory of epistemic outcome evaluation, but here's a quick and dirty proposal. Suppose that inquiry path A is just as good as inquiry path B with respect to promoting epistemic value at the end of inquiry, but B is more costly than A in terms of time (and/or other resources). We might then say that with respect to epistemic outcome evaluation there is no difference between them. Yet, from the point of view of *practical* rationality, S should, other things being equal, choose A over B, as this will free up resources for other inquiries, which in turn can result in more epistemic value in the bigger picture. On this proposal, allocation of resources is fundamentally a matter of practical rationality, not epistemic rationality.[17] This aligns with what is often assumed in epistemology, where worries about costs of inquiry in terms of time (and/or other resources) are set aside as irrelevant for epistemic concerns. Obviously, however, these issues are frequently intertwined. Inquiries may be nested and interdependent in such a way that selecting a cheaper path in one inquiry may free up resources to do what is epistemically rational in another. One may maximize one's overall success with respect to inquiry by selecting the least costly inquiry paths along the way. In this derivative manner, going for the less costly inquiry paths may be epistemically rational in the sense that it furthers the achievement of your overall epistemic goals.

Now, in a roundabout way this issue of overarching epistemic success leads us back to where we started our discussion of epistemic consequentialism, viz., trade-off cases. Clearly trade-offs can happen *internally* to inquiry, e.g., by adopting epistemically unjustified beliefs for the sake of epistemic gains at the termination of inquiry, but trade-offs can also occur between inquiries. Typically, success with inquiry I_1 does not make some unrelated inquiry I_2 less successful, except in the sense that time and resources spent on I_1 cannot be used for I_2. Often inquiries are related, such that success with I_1 furthers progress with I_2, or may even be a condition of success regarding I_2, but things are not necessarily so neat. Suppose a ***Truth Demon*** makes it the case that if you get maximum epistemic value relative to a target proposition of inquiry, then in return you'll get a million false beliefs about unrelated matters. How, from the point of view of epistemic outcome evaluation, should you choose between narrow perfection and overall disaster?[18]

[17] See (Thorstad 2022) for a recent defense of *epistemic nihilism about inquiry*, i.e., the view that there are no genuinely *epistemic* norms of inquiry.

[18] This distinction bears resemblance to what Thomas Kelly calls *narrow* and *wide* goals (Kelly 2003), e.g., wanting a true belief about the location of a certain bus stop is a fairly narrow goal to which only a very limited range of information is relevant, whereas wanting more true beliefs about the world in general is a very wide goal.

One response would be to insist that epistemic outcome evaluations are always relative to a specific inquiry. On this view, there would be no answer to what you should do in the sort of situation the truth demon puts you in. Relative to the target proposition of inquiry you should go ahead, but relative to all the false beliefs the truth demon will impute on you, you shouldn't. Yet, there is no answer to what you should do *simpliciter*, i.e., what you should do independently of your target of inquiry. There is then, a sense in which consequentialism about epistemic outcome evaluation cannot adjudicate in trade-offs between inquiries.

At first glance, this might seem deeply unsatisfactory. Though it might be a price worth paying considering the alternatives, e.g., expand the scope of epistemic outcome evaluation to cover inquiries in whole lifetimes, or even the entire effort of humanity as a whole. Such views would have extremely counterintuitive implications. To illustrate, it wouldn't—in terms of outcome evaluation—be epistemically right for me to go online in order to find out what a trip to Paris costs if I could instead have taken some inquiry-step that would generate more epistemic value in the long run (e.g., by reading two or three random Wikipedia-articles). And it gets worse. If we don't relativize epistemic outcome evaluations to specific inquiries defined by target propositions, we might get that no inquiry-step could be epistemically right unless it were part of that sequence of steps taken, over my whole life, that would result in my maximizing the epistemic value of all the beliefs I end up having in that life. Or worse yet: no single step could be right unless it's a part of that sequence of inquiry-steps which maximizes the epistemic value of all the beliefs that every rational being will have over the whole course of the existence of the universe.[19]

3 The Truth-Fairy Revisited

To finish the paper let's briefly revisit the intra-inquiry type of trade-off case presented at the beginning (Section 1). One reason for returning to the **Truth-Fairy** is that it has attracted a lot of attention in the literature on epistemic consequentialism. So, for this reason, it is worth showing how our envisioned version of epistemic consequentialism would handle the case. Recall the specifics of the **Truth-Fairy**:

> Suppose a truth-fairy exists who guarantees that, if I believe that Lake Constance is filled with milk—a proposition I have many good reasons to think is false and that is,

[19] For the analogous problems in ethics, see for example (Anderson 1995, p. 84)—she cites (Griffin 1986, p. 34) and (Slote 1985, p. 103).

> in fact, false—then she will arrange things in a way that all the other (and logically independent) beliefs I entertain now and in the future will be true. Given these circumstances, should I believe that Lake Constance is filled with milk? Would that belief be epistemically justified? (Briesen 2016, p. 288)

Applying epistemic consequentialism to the case, one could say that, while the unsupported (false) belief put forward by the truth-fairy is epistemically unjustified due to its lack of epistemic backing, it is nonetheless epistemically right in a certain sense to adopt the belief: as the case is specified adopting this particular belief—that <Lake Constance is filled with milk>—will bring about an overwhelmingly good epistemic outcome. Recall, as we mentioned earlier, that applying epistemic consequentialism as a criterion of rightness to individual beliefs adopted in inquiry is tricky since individual beliefs or inquiry-steps normally don't have determinate consequences independently of other beliefs or inquiry-steps. This was a reason to focus on a consequentialist outcome evaluation of (combinations of) decision procedures for beliefs rather than a consequentialist outcome evaluation of beliefs. By contrast, the **Truth-Fairy** stipulates sufficiently determinate outcomes of adopting just one particular belief. Imagine now a decision procedure for beliefs instructing agents to adopt one or a few false beliefs in cases where a truth-fairy guarantees this to result in a massive amount of unrelated true beliefs or successful inquiries. Applying epistemic consequentialism to decision procedures gives us that this is the right decision procedure to use, given that no other (combination of) available decision procedure(s) is better. Of course, in the actual world truth-fairy cases don't occur, and a decision procedure taking its cues from a truth-fairy doesn't work.

However, as we said earlier, inquiries with a similar structure requiring a trade-off exist, and also here we can focus on decision procedures. In **Surveying Literature** the best available decision procedure, i.e., reading through a certain body of text, comes with the cost of adopting a number of false beliefs during inquiry. Yet, this procedure is epistemically right given our consequentialist outcome evaluation, on the assumption that there is no other way for the agent to bring about an outcome with a better epistemic value regarding the target proposition at the end of inquiry. In **Unjustifiedly Believing a Theory** an epistemically unjustified belief in a controversial theory is a psychological precondition for obtaining knowledge about the target proposition of inquiry. Given that this unjustified belief is part of the best combination of available decision procedures, it is a cost worth paying, when assessed in terms of our consequentialist outcome evaluation. Similarly, in **Joint Inquiry**, a strategic distribution of cognitive efforts leads our detectives to take on an investigation

of a suspect that would have seemed irrational from a purely individualistic backing perspective, and yet—according to our consequentialist outcome evaluation—their division of cognitive labor is epistemically right given the social setting and mutual goal of inquiry (as long as their division of labor is part of a maximizing combination of available decision procedures). Finally, in **Evidence Selection**, our consequentialist criterion allows a medical doctor to deliberately neglect good pieces of evidence because this is part of the best combination of available decision procedures at the time.

4 Conclusion

Let's sum up. In this paper we have tried to reorient the debate about epistemic consequentialism. Instead of focusing on standard theories of epistemic justification, their (dis)similarity with consequentialist theories in ethics, and their vulnerability to certain trade-offs, we have focused on epistemic consequentialism seen as a metatheoretic framework of inquiry. We have initially distinguished between two different kinds of epistemic evaluation, viz., backing evaluation and outcome evaluation, and eventually outlined and discussed a version of epistemic consequentialism that seems plausible with respect to the latter. We have suggested viewing epistemic consequentialism as a higher-order criterion vis-à-vis the combinations of decision procedures we ought to adopt in inquiry, and we have tested the verdict of our consequentialist view in the widely discussed **Truth-Fairy** case.[20]

References

Andersen, F.J. and Kappel, K. (2020). Process reliabilism, prime numbers and the generality problem. *Logos & Episteme*, 11(2), 231–236.

Anderson, E. (1995). *Value in ethics and economics*. Harvard University Press.

Ahlstrom-Vij, K. and Dunn, J. (2014). A defence of epistemic consequentialism. *Philosophical Quarterly*, 64(257), 541–551.

Ahlstrom-Vij, K. and Dunn, J. (2017). Is reliabilism a form of consequentialism. *American Philosophical Quarterly*, 54 (2), 183–194.

Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, Epistemology (1988), 257–299.

[20] We are extremely thankful to the editor of the *Asian Journal of Philosophy* and three anonymous reviewers for several rounds of constructive feedback during our writing of this paper.

Berker, S. (2013a). Epistemic teleology and the separateness of propositions. *Philosophical Review*, 122(3), 337–393.

Berker, S. (2013b). The rejection of epistemic consequentialism. *Philosophical Issues*, 23 (Epistemic Agency), 363–387.

Berker, S. (2015). Reply to Goldman: Cutting up the one to save the five in epistemology. *Episteme*, 12, 145–153.

Brown, C. (2011). Consequentialize This. *Ethics*, 121 (4), 749–771.

Briesen, J. (2016). Epistemic consequentialism: Its relation to ethical consequentialism and the truth-indication principle. In M. Grajner & P. Schmechtig (eds.), *Epistemic reasons, norms, and goals*, pp. 277-306. Berlin/New-York: de Guyter.

Carlson, E. (1995). *Consequentialism reconsidered*. Dordrecht: Kluwer. Academic Publishers.

Carr, J. R. (2017). Epistemic utility theory and the aim of belief. *Philosophy and Phenomenological Research*, 95(3), 511–534.

Conee, E. (1987). Evident, but rationally unacceptable. *Australasian Journal of Philosophy*, 65, 316–326.

Fantl, J. (2021). Fake News vs. Echo Chambers. *Social Epistemology*.

Feldman, R. & Conee, E. (1985). Evidentialism. *Philosophical Studies*, 48,15–34.

Feldman, R. (2001). Voluntary belief and epistemic evaluation. In Matthias Steup (ed.), *Knowledge, truth, and duty*, pp. 77–92. Oxford: Oxford UP.

Fumerton, R. (1995). *Metaepistemology and skepticism*. Lanham, MD: Rowman & Littlefield.

Friedman, J. (2019). Teleological Epistemology. *Philosophical Studies*,176, 673–691.

Friedman, J. (2020). The Epistemic and the Zetetic. *Philosophical Review*, 129(4),501–536.

Goldman, A. (1979). What is Justified Belief?. In: Pappas G.S. (ed.) *Justification and Knowledge*. Philosophical Studies Series in Philosophy, vol 17. Springer, Dordrecht.

Goldman, A. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.

Goldman, A. (1999). *Knowledge in a social world*. Oxford University Press.

Goldman, A. (2015). Reliabilism, veritism, and epistemic consequentialism. *Episteme*, 12(2), 131–143.

Greaves, H. (2013). Epistemic decision theory. *Mind*, 122(488), 915–952.

Griffin, J. (1986). *Well-being: Its meaning, measurement and moral importance*. Oxford: Clarendon Press.

Hallsson, B. G. and Kappel, K. (2018). Disagreement and the division of epistemic labor. Synthese, 197, 2823–2847.

Hare, R.M. (1981). *Moral Thinking: Its Levels, Method, and Point*. Oxford University Press.

Howard-Snyder, F. and Norcross, A. (1993). A consequentialist case for rejecting the right. *Journal of Philosophical Research*, 18, 109–125.

Jenkins, C. S. (2007). Entitlement and Rationality. *Synthese*, 157(1), 25–45.

Kagan, S. (1998). *Normative Ethics*. Boulder: Westview Press.

Kahan, D. M., and Corbin, J. C. (2016). A note on the perverse effects of actively open-minded thinking on climate-change polarization. *Research & Politics*, 3(4).

Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, 66(3), 612–640.

Kitcher, P. (1990). The Division of Cognitive Labor. *Journal of Philosophy*, 87, 5–22.

Littlejohn, C. (2012). *Justification and the truth-connection*. Cambridge University Press.

Littlejohn, C. (2018). The right in the good: A defense of teleological non-consequentialism. In Kristoffer Ahlstrom-Vij & Jeffrey Dunn (eds.), *epistemic consequentialism*, pp. 23–47. Oxford University Press.

Nagel, T. (1989). *The view from nowhere*. Oxford University Press.

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Scott, J. (2023). Should epistemic instrumentalists be more social? *Synthese*, 201(4).

Slote, M. (1984). Satisficing Consequentialism. *Proceedings of the Aristotelian Society*, 58, 139–163.

Slote, M. (1985). *Common-sense morality and consequentialism*. Routledge.

Talbot, B. (2014). Truth promoting non-evidential reasons for belief. *Philosophical Studies*,168(3), 599–618.

Thorstad, D. (2022). There are no epistemic norms of inquiry. *Synthese*, 200(5).