Epistemic Consequentialism: Its Relation to Ethical Consequentialism and the Truth-Indication Principle

Jochen Briesen, University of Konstanz

(jochen.briesen@uni-konstanz.de)

Abstract. Consequentialist positions in philosophy spell out normative notions by recourse to certain aims. *Hedonistic* versions of *ethical consequentialism* spell out what is *morally* right/justified via recourse to the aim of increasing pleasure and decreasing pain. *Veritistic* versions of *epistemic consequentialism* spell out what is *epistemically* right/justified via recourse to the aim of increasing the number of true beliefs and decreasing the number of false ones. Even though these theories are in many respects structurally analogous, there are also interesting disanalogies. For example, popular versions of epistemic consequentialism implicitly endorse the truth-indication principle (which claims that a belief is epistemically justified only if there are factors indicating that the belief *itself* is true), whereas popular versions of ethical consequentialism do not subscribe to an analogous pleasure-indicating principle (which claims that an act is morally justified only if there are factors indicating that performing the act *itself* is pleasurable). In a first step I will argue that this difference rests on the fact that plausible versions of *epistemic consequentialism* have to meet certain constraints, which versions of *ethical consequentilialism* do not have to satisfy. As these constraints can be easily met by incorporating the truth-indication principle, epistemic consequentialists tend to subscribe to it. In a second step I will investigate whether the identified constraints can also be met independently of the truth-indication principle. Are there plausible versions of veritistic epistemic consequentialism that reject the principle, thereby allowing that some beliefs can be epistemically justified even though no factors speak in favor of their truth? Building on ideas put forward by Ludwig Wittgenstein, Crispin Wright, and others, I will answer this question affirmatively.

1. Introduction

The terms "justification" and "rationality" refer to many different relations and phenomena. Many epistemologists agree that the best way to specify *epistemic* justification/rationality and thereby differentiate it from other forms of justification/rationality is the following: *Epistemic* justification/rationality, unlike *moral* or *pragmatic* justification/rationality, is directed at an epistemic aim, whereas other forms of justification/rationality are directed at other aims. Thereby, many epistemologists adopt a consequentialist position in the sense that they spell out the normative notions of "justification" or "rationality" by recourse to certain epistemic aims.[1]

The main motivation for adopting such a consequentialist position in epistemology is its explanatory strength. Consider the following case:

> *Criminal child.* The child of person *S* is accused of a terrible crime based on very good evidence. Even though *S* knows about all the evidence and can neither undermine nor override it in any way, *S* nonetheless holds onto the belief that the child is innocent.

From a pretheoretical perspective, it seems plausible that *S*'s belief can be justified/rational in a certain sense, whereas in another sense the belief seems clearly unjustified/irrational. This can be easily explained within the above-mentioned epistemic consequentialist framework: *S*'s belief might be *pragmatically* justified because entertaining the belief that one's child is innocent might be conducive to a person's *practical* aim; maybe the person aims at living a happy life and would not be able to achieve that aim if she gave up the belief. However, *S*'s belief is obviously not *epistemically* justified/rational because entertaining the belief that the child is innocent does not seem to be conducive to an *epistemic* aim, for example the aim of increasing the number of true beliefs and decreasing the number of false ones. In the light of cases such as this, many epistemologists take truth-conduciveness to be a defining feature of

---

[1] One might claim that the given characteriziation only specifies what we might call a "teleological" and not necessarily a consequentialist position. However, in the context of this paper I will not differentiate between the two. I will use the term "consequentialism" to refer to all positions that spell out normative notions via recourse to the aim of achieving something of final value.

epistemic justification/rationality and thereby adopt an epistemic consequentialist position.

Even though the whole epistemic consequentialist framework has been recently attacked by Selim Berker (2013a, 2013b), I will accept the general framework in what follows. Berker's considerations do not seem to constitute devastating objections to the general framework, and the necessity of explaining the differences among various forms of justification/rationality gives us very good reason to adopt such a framework.[2]

Moreover, even though I agree that different important epistemic values have to be distinguished, for the purpose of this paper, I accept that one very important epistemic value is truth, and one very important epistemic disvalue is falsity.[3] Thus, I will accept that increasing the number of true beliefs while avoiding false beliefs can be considered an important epistemic goal by recourse to which some important epistemic notions can be specified. (I will mostly concentrate on the notion of epistemic justification.) Hence, I will accept a *veritistic* consequentialist framework.

What I am interested in is the question of how such a framework should be spelled out in more detail. What are the constraints for a convincing veritistic consequentialist position? In what way can these constraints be met? To answer these questions, it is helpful first to consider the putative analogy between epistemic and ethical consequentialism. Even though epistemic and ethical variants of consequentialism are in many respects structurally analogous, there are also important disanalogies. Some of these disanalogies will eventually help to identify constraints that epistemic versions of consequentialism have to meet. Identifying these constraints will in turn allow consideration of different ways in which the constraints might be satisfied. In following this general strategy, the paper is organized as follows:

In section 2, I will first specify structural analogies between epistemic and ethical variants of consequentialism. After highlighting the analogies, I will, in section 3, identify an important disanalogy with respect to these positions. In this context, I will introduce the following "truth-indication principle" (henceforth (TI)).

---

[2] I will hint at some reasons why Berker's objections do not seriously undermine the epistemic consequentialist framework below (see fn. 22). It should also be noted that Berker does not take the reasons that speak *in favor* of such a framework seriously into account.

[3] For interesting discussions of various epistemic values, see, e.g., Haddock et. al 2009; Grimm 2008; Kvanvig 2005.

(TI) For all subjects *S* and propositions *p*: *S*'s belief that *p* is epistemically justified only if there are factors indicating that the belief *itself* is true.

Even though this principle is rarely explicitly mentioned, I will argue that popular consequentialist positions in epistemology (internalist and externalist versions alike) implicitly endorse (TI). However, no analogue principle (for example, a pleasure-indication principle) can be found within ethical consequentialist positions. Thus, from a certain perspective, an important difference between epistemic and ethical consequentialism amounts to the fact that epistemic versions of consequentialism seem to be committed to (TI), whereas no analogous principle is warranted in the ethical domain. The reason epistemic consequentialists tend to incorporate (TI) into their framework is that, by incorporating (TI), they can meet certain constraints that every plausible version of epistemic consequentialism has to satisfy. In the course of section 3, I will identify these constraints in detail.

In section 4, I will consider whether the identified constraints can only be met by accepting (TI) or also be met differently. In other words: Are there plausible versions of epistemic consequentialism that reject the truth-indication principle (TI)? If there were, then the analogy between epistemic and ethical consequentialism would, in a certain sense, be even stronger than it seems at first. By specifying and modifying some ideas famously introduced by Ludwig Wittgenstein (1969) and further developed by Crispin Wright (2004, 2014) and others, I will spell out a promising epistemic consequentialist position that meets the identified constraints but is nevertheless not committed to (TI).

2. The analogy between ethical and epistemic consequentialism

Consequentialist positions in philosophy spell out normative notions by making recourse to the goal of achieving states of affairs that have final value. A state of affairs has final value if and only if it has value as end in itself (in contrast to having instrumental value as a means to some other valuable end). To keep things simple in terms of theories of value, I will concentrate on a simple version of *hedonism* in ethics, which claims that experiencing pleasure has final value and experiencing pain has final disvalue, and a simple version of *veritism* in epistemology, which claims that having

true beliefs has final *epistemic* value and having false beliefs has final *epistemic* disvalue.[4]

By accepting these final values, consequentialists in ethics and epistemology can formulate important aims, namely the aims of achieving the states of affairs of final value in their respective domains. In what follows, I will call these aims "ultimate aims". For *hedonistic consequentialists*, the ultimate aim in the ethical domain consists in improving the overall balance of pleasurable versus painful experiences (by increasing the number of pleasurable experiences and decreasing the number of painful ones). For *veritistic consequentialists,* the ultimate aim in the epistemic domain consists in improving the overall balance of (interesting) true beliefs over false ones (by inreasing the number of true and decreasing the number of false beliefs).[5][6]

By making recourse to these aims, those taking consequentialist positions in ethics and epistemology can put forward normative evaluations. For example, by referring to the specified aims, hedonistic consequentialists in ethics can evaluate *acts* as being *morally* justified*,* and veritistic consequentialists in epistemology can evaluate *beliefs* as being *epistemically* justified:

> *(MJ) Moral Justification of Acts.* For all subjects *S* and acts *A*: *S*'s action *A* is morally justified only if performing *A* is conducive to the ultimate aim of improving the overall balance of pleasurable experiences over painful experiences (to a certain degree).[7]

---

[4] Note that according to this characterization veritists claim that true beliefs have *final epistemic* value and not *epistemic final* value. Thus, veritist consequentialists merely claim that true beliefs have final value in the *epistemic domain*. I will follow the common practice of taking this formulation to be indifferent with respect to the question whether true beliefs have this value only in virtue of being conducive to some other value outside the epistemic domain.

[5] It seems plausible to formulate the epistemic aim as improving the balance of *interesting* true beliefs over false ones. Otherwise, it would be epistemically appropriate to concentrate exclusively on beliefs that can be easily verified. Examples include beliefs concerning the number of grains of sand in our shoes and beliefs concerning the number of coffee beans in jars. By specifying the epistemic aim as being directed at *interesting* true beliefs, epistemic consequentialists can explain why this kind of behaviour would be epistemically inappropriate. Why I nonetheless put the term "interesting" in parentheses in the formulation of the ultimate epistemic aim in the main body of the text will be explained below (see. fn. 8).

[6] Note the double character of this aim, it can be achieved either by increasing the number of (interesting) true beliefs or by decreasing the number of false ones. That the epistemic truth-aim has this kind of double character rests on insights by William James (cf. James 1897:18).

[7] These kinds of evaluations in the context of ethical consequentialism are variable with respect to two important parameters. First, both acts and other things can be morally evaluated: e.g. motives, maxims, institutions, and even beliefs. Second, acts and other things can be evaluated as being morally justified as

(EJ) Epistemic Justification of Belief. For all subjects S and propositions p: S's belief that p is epistemically justified only if having the belief that p is conducive to the ultimate epistemic aim of improving the overall balance of (interesting)[8] true beliefs over false beliefs (to a certain degree).[9]

Different versions of these positions can be characterized by further specifying what it means to be "conducive" to the respective aims.

Let us first consider some important variations within ethical hedonistic consequentialism.[10] Hedonistic consequentialists agree that, for an act A to be conducive to the specified aim, there must be some feature of A that will be accountable for a greater balance of pleasurable over painful experiences. However, different views take different features of A to be relevant. *Act*-consequentialists, for example, claim that the relevant feature is the causal chain of events set off by A. If this chain of events leads to a certain improvement of the balance of pleasure over pain, then the act is morally justified. *Motive*- or *rule*-consequentialists, on the other hand, think that the relevant feature is not the causal chain of events set off by A, but rather the motive-set M or the set of rules R according to which S performed A. If acting according to M (when possessed in appropriate circumstances) or according to R (when properly internalized by a sufficiently large group of the population) leads in the long run to a certain improvement in the balance of pleasure over pain, then the act in accordance with M or R is morally justified.

---

well as being morally right, being morally required, being morally permitted, etc. For more variable parameters, see fn. 11.

[8] Even though the epistemic aim should be specified as being directed at *interesting* true beliefs, we should allow that uninteresting beliefs can be justified as well. Take the belief, for example, that there are 321 grains of sand in my shoe. Let us suppose that holding this uninteresting belief cannot have any positive consequences with respect to increasing the number of *interesting* true beliefs. Nevertheless, we should allow that this belief could be justified. This is why I put the term "interesting" in parentheses in the formulation of (EJ). Even though the ultimate epistemic aim should be spelled out as being directed at increasing the number of *interesting* true beliefs, in our epistemic evaluation of beliefs sometimes the more general aim of increasing the number of true beliefs *simpliciter* seems relevant.

[9] Just as evaluations in the context of *ethical* consequentialism are variable, so are the evaluations in the context of *epistemic* consequentialism. First, not only beliefs but also other things can be epistemically evaluated: e.g. belief-sets, theories, belief-forming processes, and even acts. Second, beliefs or other things can not only be evaluated as being epistemically justified, they can also be evaluated as being epistemically rational, epistemically correct, epistemically required, epistemically permitted, etc. See also fn. 11.

[10] For a general overview of different versions of consequentialism in ethics, see Sinnot-Armstrong 2014.

It is important to note that, in the light of motive- or rule-consequentialism, act *A* is conducive to the specified aim in an indirect way. This is why these variants of consequentialism are sometimes subsumed under the term "indirect consequentialism". In a certain sense, the act itself is not supposed to be conducive to the specified aim in order to be morally justified, but rather the motive set *M* or the rules *R* on which act *A* is based. This allows that, in individual cases, acting in accordance to *M* or *R* can even have negative effects with respect to the specified aim without thereby necessarily being unjustified. As long as acting in accordance to *M* or *R* will lead to a greater balance of pleasure over pain *in the long run,* an individual act can even have negative consequences with respect to the specified aim and still be morally justified.

Besides identifying the feature of act *A* that is accountable for the act's conduciveness towards the specified aim and thereby specifying in which way the act has to be conducive to the aim, hedonistic consequentialists have to decide whether the identified features of *A* have to improve the *actual* balance of pleasure over pain or whether it is enough that subject *S* reasonably *believes* that the ratio is improved. In other words: Are the features such that they improve the balance of pleasure over pain from a third- or a first-person perspective? *Objective act*-consequentialists hold that what is morally relevant is whether the *actual* chain of events set off by *A* leads to a greater balance of pleasure over pain. *Subjective act*-consequentialists, on the other hand, claim that what is morally relevant is not the actual chain of events, but rather the chain of events *intended* or *expected* by the agent. A similar distinction can be drawn within *motive-* or *rule*-consequentialism in ethics. Is it morally relevant whether acting according to motive-set *M* or rules *R actually* improves the balance of pleasure over pain in the long run, or is it relevant whether the subject *intends* or *reasonably expects* that this is the case?

Even though there are more variable parameters within hedonistic consequentialism, the parameters introduced so far will suffice to spell out the analogies and disanalogies between ethical and epistemic consequentialist theories in which I am interested. So I will turn to epistemic variants of consequentialism. [11]

---

[11] There are two more variable parameters within hedonistic consequentialism that should at least be mentioned. The first concerns the question of to what degree the balance of pleasure over pain has to be improved, i.e. is a satisficing or a maximizing version of hedonistic consequentialism preferred? Even though an analogous question could be raised in the epistemological realm, for the purposes of this paper this analogy is negligible. The second parameter concerns whether an egoistic or a non-egoistic form of hedonistic consequentialism should be preferred. In egoistic versions, only the ratio of the *subject's*

Veritistic consequentialists in epistemology agree that, for a belief B to be conducive to the epistemic aim of increasing the number of true beliefs and decreasing the number of false ones, there must be some feature of B that will be accountable for a greater balance of true over false beliefs. However, as different ethical theories disagree with respect to the relevant features of act A, different versions of veritistic consequentialism in epistemology take different features of belief B to be relevant. *Evidentialist* versions of consequentialism, for example, think that the relevant feature is the body of evidence E on which B is based. If believing in accordance with E leads to a greater balance of true over false beliefs, then the belief appropriately based on E is justified.[12] Conversely, process-reliabilists think that the relevant feature is not the body of evidence available to S, but rather the belief-forming process P that caused S to entertain the belief B (where this process does not have to include evidence). If P is sufficiently reliable—if it leads to true beliefs most of the time—then the belief based on P is epistemically justified.[13]

Furthermore, veritistic consequentialists in epistemology not only disagree about the features accountable for a better ratio of true over false beliefs, they also disagree whether the identified features have to *actually* improve the balance in question or whether it is enough that the subject S *thinks* or *believes* that the balance is improved. In other words: Are the features such that they improve the balance of true over false beliefs from a third- or a first-person perspective or both? Pure objective epistemic consequentialists vote for the first option, pure subjective epistemic consequentialists vote for the second option, and mixed positions claim that the balance has to be improved both from a first- and third-person perspective.[14] [15]

---

pleasurable over painful experiences is relevant, whereas in non-egoistic versions the experiences of *other people* have to be taken into account. For obvious reasons, epistemic versions of consequentialism are more closely related to egoistic versions of ethical consequentialism than to non-egoistic versions.

[12] For views coming close to such a position, see Alston 1989; BonJour 1985, 2003; Conee 1992; Feldman 2000; Foley 1987, 1993. These philosophers agree that in order for a belief to be justified it has to be conducive to the epistemic truth-aim, where what is accountable for the conduciveness to the specified aim is the subject's available body of evidence on which she is basing the belief. (Please note, however, that all these philosophers spell out the epistemic truth-aim slightly differently.)

[13] This very influential view has been suggested, developed, and defended by Alvin Goldman. He thinks that in order for a belief to be justified it has to be conducive to the epistemic truth-aim, where what is accountable for the conduciveness to the specified aim is the process leading to the belief (cf. Goldman 1979, 1986, 2001).

[14] For a view coming close to a *pure* subjective position, see Lehrer 1974; a good example of a mixed position can be found in Alston 1989, 2005.

[15] In a certain usage of the terms "epistemic internalism" and "epistemic externalism", the distinction between *subjective* and *objective* consequentialism amounts to the difference between epistemic

Please note that evidentialist and process-reliabilist versions of veritistic consequentialism can both be classified as forms of *indirect* consequentialism. In a certain sense, not the belief itself is supposed to be conducive to the epistemic aim in order to be justified but rather the body of evidence *E* or the belief-forming process *P* on which *B* is based.

This is not an accident. Epistemic consequentialism tends to be spelled out as an instance of indirect consequentialism because it is pretheoretically plausible to assume that false beliefs can be epistemically justified. It seems plausible to suppose that, for example, in certain circumstances, my belief that my bike has been stolen is justified, without that belief's being true. Perhaps, maybe unbeknownst to me, my neighbour borrowed the bike.[16]

Subjective epistemic consequentialists can easily account for this possibility. However, if *objective* epistemic consequentialists want to allow that a false belief can be epistemically justified, then they have to claim that, in a certain sense, it is not the justified belief itself that has to be conducive to the specified epistemic aim but rather something else, such as the belief-forming process *P* or the body of evidence *E* on which *B* is based. This allows for the possibility that, in individual cases, a belief that is caused by a reliable process *P* or that is appropriately based on evidence *E* can be false and can thereby have negative effects with respect to the specified epistemic aim, without that belief's thereby necessarily being unjustified. As long as believing in accordance to *P* or *E* will lead to a greater balance of true over false beliefs *in the long run*, the false belief in question can still be epistemically justified. It is interesting to note that, as *indirect* versions of consequentialism, epistemic consequentialist theories are more closely related to ethical motive- or rule-consequentialism than to ethical act-consequentialism.

Up to this point, the similarities between ethical and epistemic forms of consequentialism are pretty straightforward. Both have to specify due to which features and in what way a morally/epistemically justified act/belief is conducive to the respective aims. This decision will settle whether a *direct* or an *indirect* variant of

---

internalism and epistemic *externalism*. However, because the latter terms are used differently in the literature it is better to stick to the terms "subjective epistemic consequentialism" and "objective epistemic consequentialism" with their stipulated meanings.

[16] Even though most epistemologists think that there can be justified false beliefs, not everyone agrees. For a defense of the opposite view that justification is factive, see Littlejohn 2012.

consequentialism is accepted. Consequentialists of both camps also have to specify whether the identified features are supposed to be conducive to the respective aims from a first- or third-person perspective or both. This decision will settle whether a *subjective* or an *objective* form of consequentialism (or a mixed position) is accepted.

3. An important disanalogy between ethical an epistemic consequentialism

There is a difference in the way we set things up between ethical and epistemic variants of consequentialism that catches one's eye. The versions of epistemic consequentialism characterized above are all committed to the view that a belief is epistemically justified partially in virtue of factors speaking in favor of the truth of the justified belief itself, be it from a first- or a third-person perspective or both. Thus, the characterized positions are all committed to the following truth-indication principle (hereafter, (TI)):

> (TI) For all subjects S and propositions p: S's belief that p is justified only if there are factors indicating that the belief *itself* is true.[17]

The epistemic consequentialist positions characterized in section 2 claim that the justified belief is conducive to the global epistemic aim of improving the balance of true over false beliefs because there are some factors—either the body of available evidence E or the reliability of the belief-forming process P, etc.—that (to a certain extent) indicate—either from a first- or third-person perspective or both—that the belief itself is true. This is why holding justified beliefs will ultimately lead to a greater balance of true over false beliefs either from a first- or third-person perspective or both. Thus, in a certain sense, holding justified beliefs has positive epistemic consequences globally only because a justified belief has positive epistemic consequences locally. Holding justified beliefs has the positive *global* consequence that a greater balance of true over false beliefs is achieved in the long run. However, this positive global consequence is

---

[17] It is not easy to spell out in detail what it means to say that X partially in virtue of Y. I assume it implies that whenever X is present, Y is present. Thus, it implies that X only if Y. This is why a position that claims that a belief is epistemically justified partially in virtue of factors indicating the truth of the justified belief itself is committed to (TI).

due to the positive *local* consequence that the justified belief *itself* is likely to be true—again, either from a first-person perspective, a third person perspective or both. [18] [19]

In contrast, the characterized versions of hedonistic consequentialism are not committed to an analogous pleasure-indication principle. What would such a principle look like? If we simply substitute "epistemically justified" in (TI) by "morally justified" and any instance of "belief" by "act" and "truth/true" by "pleasure/pleasurable", we yield the following pleasure-indication principle (PI):

> (PI) For all subjects *S* and acts *A*: *S*'s act *A* is morally justified only if there are some factors indicating that the act *itself* is pleasurable.

In contrast to (TI), (PI) is quite implausible. First of all, in the way (PI) is formulated, it is not even intelligible because, although a belief itself can be true or false, an act itself cannot be pleasurable or painful. Only the *experiences* a subject has in *performing* an act can be pleasurable or painful. Thus, the pleasure-indication principle has to be formulated slightly differently:

> (PI)* For all subjects *S* and acts *A*: *S*'s act *A* is morally justified only if there are some factors indicating that *S*'s *experiences in performing* act *A itself* are pleasurable.

In this form, the pleasure-indication principle is intelligible, but none of the versions of hedonistic consequentialism characterized in section 2 is committed to (PI)*. Hedonistic consequentialists do not subscribe to (PI)* because the principle seems clearly false. It is obviously not a necessary condition for an action to count as morally justified that the

---

[18] In a certain strand of Bayesian epistemology, known as epistemic utility theory, global and local epistemic aims are standardly differentiated (cf., e.g., Leitgeb and Pettigrew 2010). I am suggesting that an analogous distinction should be drawn within a binary model of belief as well. The global aim concerns our whole belief-set and consists in improving the balance of true over false beliefs within this set. The local aim concerns individual beliefs and can roughly be characterized as: Believe that *p* if and only if *p* is true. The idea is that by subscribing to (TI), achievements with respect to the global aim are in a certain sense mediated by achievements with respect to the local aim.

[19] In the following quote, Carrie Ichikawa Jenkins claims that mediating positive global consequences via positive local consequences should be considered a constraint for plausible versions of epistemic consequentialism: "The basic thought is roughly that we shouldn't be able to establish that accepting S is epistemically rational just by establishing that such acceptance has optimal epistemic consequences *overall*, and without needing to make reference to *particular* epistemic goals we have with respect to S" (Jenkins 2007: 36).

agent experiences pleasure during the performance of this *very action*. Assume I am an incredibly penurious person, so that in donating money I literally experience pain. Is my action of donating money thereby condemned to be morally unjustified? Surely not. Thus, ethical consequentialists have very good reason to reject (PI)*.[20]

In contrast, incorporating (TI) into the *epistemic* consequentialist framework seems like an attractive move because, by subscribing to (TI), consequentialists can handle certain well-known and obvious counter-examples to crude versions of epistemic consequentialism.

> *Grant-Seeking.* Person *X*, a brilliant virologist and an agnostic, is seeking to obtain a research grant from a religious organization. If she receives the funding, she will be able to continue her research, which will eventually allow her to form a great many true beliefs and to revise many previously held false beliefs about matters of great importance. However, the funding organization will only give grants to persons who believe in the existence of God. Given these circumstances, should *X* form the belief that God exists? Would such a belief be epistemically justified (cf. Fumerton 2001: 55)?

Prima facie, this case seems to be a counter-example to epistemic consequentialism. After all, *X*'s belief in the existence of God would clearly be conducive to the aim of improving the balance of interesting true beliefs over false ones. However, pretheoretically it is quite implausible to judge the belief in question as justified.

By incorporating (TI) into the consequentialist framework, epistemic consequentialists can easily dismiss this case. Even though *X*'s belief concerning the existence of God is in a certain sense conducive to the epistemic aim, it is not conducive to that aim in the right kind of way, i.e. in the way specified by (TI). There are no factors indicating that the proposition in question is true—in the example, *X* neither has available evidence for the existence of God, nor would *X's* belief forming method be

---

[20] One might think that (PI)* is only unattractive for someone rejecting egoistic versions of hedonistic consequentialism. However, that's not true. As long as *S*'s act *A* will eventually lead to a greater balance of *S*'s pleasurable experiences over her painful experiences, even egoist versions of consequentialism will usually allow that *A* might very well be justified, even though performing *A* is accompanied by some pain for *S*.

particularly reliable. Thus, even though the belief is in a certain sense conducive to the epistemic aim, it is nevertheless epistemically unjustified.

Counter-examples like these can be found in various places in the literature.[21] Before I localize the general lessons to be learned from these cases, let me introduce another very drastic one:

> *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 (cf. Jenkins 2007: 37)?

Again, my belief that Lake Constance is filled with milk is clearly conducive to the epistemic aim. However, pretheoretically it is overwhelmingly plausible to think that this belief would be epistemically unjustified. By incorporating (TI) into the consequentialist framework, epistemic consequentialists can dismiss this case. My belief that Lake Constance is filled with milk is not truth-conducive in the right kind of way, i.e. in the way specified by (TI). [22]

---

[21] For a list of references with similar uses of similar cases, see Berker 2013a: fn. 38.

[22] It should be noted that by incorporation of (TI) not all counter-examples to epistemic consequentialism can be dismissed. Two examples discussed by Berker still pose problems, namely the prime-number case (Berker 2013a: 375) and the John-Doe case (cf. Berker 2013b: 369; for a slightly modified version of the case, see p. 376). It could be argued that in both cases, beliefs are characterized that are not only conducive to the general epistemic goal; there are also certain factors indicating that the beliefs *themselves* are true. Nonetheless, in these cases we pretheoretically judge the beliefs in question as not being epistemically justified. Thus, consequentialists have to handle these cases differently. Simply incorporating (TI) into their framework will not allow resisting the pretheoretically implausible result with respect to these cases.

However, I am unsure whether this poses an insurmountable problem for consequentialists. In the context of this paper I can only hint at ways in which consequentialists could react to these cases. First, a general remark: By incorporating (TI) into the consequentialist framework, most counter-examples put forward in the literature can be dismissed, and the remaining counter-examples do not show that there is some deep structural problem with epistemic consequentialism. Rather, they show that some specific formulations of epistemic consequentialism are problematic.

The prime-number case (Berker 2013a: 375; 2013b: 374-375), for example, only establishes a problem for *reliabilist versions* of consequentialism. And in the face of other well-known problems and counter-examples, reliabilists have to enhance and modify their accounts anyway. As far as I can see, some of these modifications will allow them to get to grips with the prime-number case as well. Berker himself mentions one such modification: The reliabilist theory could be modified such that what matters with respect to the justificatory status of a belief is not just whether it is formed by a reliable process, but

There are two general lessons to be learned from considering these two counter-examples. Epistemic consequentialists claim that the epistemic evaluation of a belief that p depends on positive consequences of entertaining that belief with respect to the specified epistemic end. However, plausible versions of epistemic consequentialism have to satisfy the following constraints.

(a) *The Means-End Constraint.* The means-end relation of holding the justified belief that p and the specified epistemic aim of achieving a greater balance of true over false beliefs should be as close and direct as possible. It should not be remote (in the sense that it holds only due to many intermediary steps) or arbitrary.

(b) *The No-Trade-Off Constraint.* Trade-offs between epistemic goods with respect to different propositions that involve guaranteed epistemic sacrifices (however small they might be) should not be allowed. It should not be allowed that one knowingly makes epistemic sacrifices with respect to a belief that p to achieve a greater epistemic good with respect to beliefs in other propositions.[23]

The importance of the first constraint is vividly illustrated by the grant-seeking case and the importance of the second by the truth-fairy example. The latter illustrates that, even in drastic cases where we only have knowingly to trade in one insignificant and uninteresting false belief to guarantee the truth of *all* other beliefs we have—beliefs that

---

also whether there is an alternative reliable belief-forming process available to the subject S that, had it been used, would have resulted in that belief's not being held (cf., Berker 2013a: fn. 55). Berker quickly dismisses this strategy for the simple reason that we can just stipulate that in the prime-number case at hand such an alternative process is *not* available to S. However, I am not sure whether this can really be reasonably stipulated. It seems right to suppose that as soon as S entertains the beliefs that n is a prime number for various n and thereby grasps the concept PRIME NUMBER, an alternative belief-forming method than the one described by Berker in his example is available to S that, had it been used, would have resulted in S's *not* believing that 7 is not a prime number.

As for the John-Doe-case, it has the same kind of structure: A person holds the belief that p, where this belief has positive epistemic consequences and there are also some factors indicating that the belief itself is true. Nevertheless, in considering the John-Doe-case we pretheoretically judge that the belief in question is unjustified. However, Berker admits that the person's belief is only obviously unjustified as long as we think that the person is *not aware* of the truth-indicative factors with respect to that belief. As soon as we suppose that she is aware of these factors, it is pretheoretically not so obvious anymore that the belief is unjustified. Thus, the case seems to rest on familiar internalistic intuitions. The case does not show that there is a general problem with the consequentialist framework, but rather that externalistic versions of such a framework are problematic.

[23] Berker 2013a argues that the most pressing problem of epistemic consequentialism is its countenance of this kind of trade-off.

might be very important and interesting to us—we pretheoretically do not countenance these kind of epistemic trade-offs.

An epistemic consequentialist can easily meet these two constraints by incorporating (TI) in the consequentialist framework. First, incorporating (TI) ensures that the positive epistemic consequence of holding a justified belief with respect to the general epistemic aim is mediated by local positive epistemic consequences with respect to the justified belief itself. According to (TI), there have to be some factors that speak in favor of the truth of the justified belief either from a first- or a third-person perspective or both. The means-end relation between holding the belief and the epistemic aim of reaching a greater balance of true over false beliefs—either from a first- or a third-person perspective or both—is thereby not arbitrary and remote. Second, by incorporating (TI) into the epistemic consequentialist framework, the specified epistemic trade-offs are prevented.[24] (TI) prevents an epistemically inferior belief from counting as epistemically justified just because many other epistemically excellent beliefs are gained by holding this epistemically inferior belief. Thus, recourse to constraints (a) and (b) helps explain why (TI) is quite plausible within the framework of epistemic consequentialism.

We can also identify additional reasons based on this explanation, besides the one already given, as to why the analogue principle (PI)* in the ethical domain is implausible. Plausible versions of ethical consequentialism do not have to meet the ethical analogues to the epistemic constraints (a) and (b). First, it is implausible to think that the means-end relation between undertaking act *A* and the general aim of achieving a greater balance of pleasure over pain has to be spelled out so as to rule out many intermediary steps (cf., constraint (a)). My act of donating money is obviously not condemned to be morally unjustified just because this act leads to a greater balance of pleasure over pain only via many intermediary steps. Second, it is also implausible to suppose that every kind of trade-off in the ethical domain has to be prevented (cf. constraint (b)). On the contrary, it seems plausible to think that, in certain circumstances, painful experiences can be compensated for by many other pleasurable experiences. For example, if I could heal every suffering patient by pinching myself and my friend in the arm, this act of pinching could very well be morally justified. At least

---

[24] Other suggestions to prevent these trade-offs are discussed in Berker 2013b: 375-376 and Jenkins 2007: 36–40.

in drastic scenarios like this, where the sacrifice with respect to the specified aim is very small (a little pain for my friend and me) and the benefits gained very high (no actual and future pain for anybody else), it seems overwhelmingly plausible that trade-offs should *not* be prevented.[25]

So far we have seen that an important disanalogy between ethical and epistemic consequentialism lies in the fact that the latter position has to satisfy constraints (a) and (b), the former does not have to meet analogous constraints. We have also seen that epistemic consequentialists can meet these constraints by incorporating (TI) into their framework, whereas ethical consequentialists are well advised not to incorporate the analogous principle (PI)*. In the next section, I will consider the question of whether constraints (a) and (b) could also be met independent of (TI). Are epistemic consequentialists committed to (TI) or are there plausible versions of epistemic consequentialism that reject (TI), but nevertheless satisfy constraints (a) and (b)? If there were such a position, then the analogy between epistemic and ethical consequentialism would in a certain sense be stronger than one might think. It is true that, in contrast to *ethical* consequentialism, *epistemic* consequentialism has to satisfy constraints (a) and (b). However, if these constraints could be met independent of (TI), then just as hedonistic consequentialism in ethics is not committed to the pleasure-indication principle (PI)*, so veritistic consequentialism in epistemology would not be committed to the truth-indication principle (TI).

4. Epistemic consequentialism without the truth-indication principle

With respect to many beliefs the truth-indication principle (TI) seems plausible. If, for example, my belief that it is raining right now is justified, then it seems plausible that this belief is justified through factors speaking in favor of its truth either from a first- or a third-person perspective or both. However, are there any exceptions to the principle?

---

[25] However, it is well known that countenancing trade-offs is also a source of serious problems of ethical consequentialism. After all, there are many cases where such trade-offs seem implausible (cf. Thomson 1976). This leads to one of the most serious challenges for consequentialist theories in ethics, namely restricting the aggregative element in the theory, so as to allow for some trade-offs but not for others. It should be clear, however, that this restriction cannot be achieved by incorporating (PI)* into the theory, because by incorporating (PI)* all trade-offs would be prevented across the board.

Is (TI) as a general principle correct?

If one rejects (TI), then one allows that there are at least *some* beliefs that are non-truth-indicatively justified, i.e. beliefs that are justified independently of factors speaking in favor of their truth. If veritistic consequentialists want to allow for such beliefs (thereby allowing for exceptions to the general principle (TI)) then they have to give satisfying answers to the following questions:[26]

(1) Which beliefs have a chance of being non-truth-indicatively justified?

(2) In which sense are non-truth-indicatively justified beliefs *epistemically* justified, i.e. in which sense are they conducive to our epistemic aim of improving the balance of (interesting) true beliefs over false ones?

(3) Is there a way to spell out the conduciveness of non-truth-indicatively justified beliefs towards the specified truth-aim that respects constraints (a) and (b)?

I will try to answer these questions based on some some ideas put forward by Ludwig Wittgenstein (1969), Michael Williams (1996: 121–134; 2001: 146–172), and most explicitly Crispin Wright (2004). These philosophers agree that beliefs in certain propositions have a special epistemic status, namely beliefs in presuppositions of scientific inquiries or cognitive projects.[27] I take this to be a promising starting point to answer questions (1)–(3) satisfactorily. Beliefs in presuppositions of cognitive projects are candidates of non-truth-indicatively justified beliefs. However, questions (2) and (3) especially can only be answered satisfactorily if it is specified what cognitive projects are and what kind of presuppositions they make.

---

[26] Note that non-veritistic epistemic consequentialists obviously have more resources to account for the existence of non-truth-indicatively justified beliefs than veritistic versions of consequentialism.

[27] Each of these philosophers can be interpreted as claiming that the special status of beliefs in presupposition of cognitive projects consists precisely in being candidates for beliefs that are justified independently of truth-indicative factors. However, in the context of this paper I will not defend this exegetical claim. It should be noted, though, that Wright's theory of entitlements is most explicit with respect to this point, even though he does not use the term "presuppositon" for the propositions in question anymore (Wright 2014, fn. 3). Wright specifies entitlement as a form of epistemic warrant that is independent of truth-indicative factors (cf. Wright 2004; 2014). However, it is unclear according to Wright's view whether the propositional attitude of *belief* is conceptually tied to truth-indicative factors, namely to reasons/evidence. Maybe it is conceptually impossible to believe *p* without having any kind of evidence (broadly construed) in favor of *p*. Therefore, Wright only claims that one can be entitled to *accept* or *trust p* where the propositional attitudes of *acceptance* and *trust* are supposed to be not as closely tied to evidence as the attitude of belief (cf. Wright 2004: 175–178). In what follows, I will ignore this complication. I will assume that it is possible to entertain the propositional attitude of belief, even though there are no truth-indicative factors that speak in favor of the truth of the belief.

It is widely agreed that scientific inquiries and cognitive projects in general can be understood as attempts to formulate and answer questions (cf., Dewey 1938: 105; Hookway 1996: 7; Wright 2014: 215; etc.). If we accept that there is a close connection between cognitive projects and questions, then we might try to analyze the structure of cognitive projects by recourse to the semantics of questions. Perhaps the structure and presuppositions of a cognitive project can be clarified by analyzing the meaning of the question that in a certain sense constitutes the project. One approach to the semantics of questions is especially helpful in this respect, namely, a certain variant of the proposition-set theory. In subsection 4.1. I will introduce this theory. Fortunately for our purposes, introducing the basics of this theory will suffice. In subsection 4.2. I will then analyze the structure of cognitive projects via recourse to the introduced theory. Finally, in subsection 4.3 I will specify a version of consequentialism that allows for non-truth-indicatively justified beliefs and answer questions (1)–(3) on the basis of the results in 4.1. and 4.2.

4.1 The proposition-set theory of questions

Within truth-functional or possible-worlds semantics the meaning of a declarative sentence p is considered to be a function from possible worlds into the set of truth-values {true, false}. Because these functions are equivalent to sets of possible worlds in which p is true, the meaning of a declarative sentence p (i.e. the proposition expressed by the sentence p) is simply considered to be the set of worlds in which p is true.

Can this general idea be transferred to interrogative sentences?[28] Obviously, an interrogative sentence *?p* is neither true nor false. Thus, its meaning cannot be identified with the set of worlds in which it is true. The basic idea guiding the proposition-set approach to interrogatives is: understanding the meaning of an interrogative sentence *?p* amounts to knowing what counts as a possible answer to it. Thus, the meaning of *?p* cannot be identified with a set of worlds in which *?p* is true but with a set of possible answers to *?p* or, more precisely, with the set of meanings of possible and congruent answers. Because answers are taken to be the meanings of declarative sentences, the

---

[28] In what follows I will concentrate on direct questions and will not consider issues concerning embedded interrogatives.

meaning of a question amounts to a set of propositions, i.e. a set of sets of worlds (cf., Hamblin 1958, 1973; Karttunen 1977; Gronendijk & Stokhof 1982, 1984, 1997).[29]

Take a look at the following interrogative sentence: "Who left the party?" For simplicity's sake, let us suppose that there is no possible world with more than two people, Maria and Joseph. Proposition-set theorists specify the meaning of "Who left the party?" as the set of congruent answer-propositions, i.e. the set of propositions expressed by the sentences: "Maria left the party", "Joseph left the party", "Maria and Joseph left the party", and "Nobody left the party".[30]

Note that the proposition expressed by the sentence "Maria left the party" does not exclude that Joseph left as well. However, the answer "Maria left the party" is in a certain sense incomplete if Joseph also left and the speaker knows that. This observation can be incorporated into the semantic framework by following Gronendijk and Stokhof's suggestion that answer-propositions form a mutually exclusive and complete partition of possible worlds (cf. Gronendijk & Stokhof 1984; 1997).[31] In such a system the meaning of "Who left the party?" divides a certain subspace of the space of possible worlds into non-overlapping cells that correspond to the propositions that *only* Maria left, that *only* Joseph left, that Maria *and* Joseph left, and that nobody left.

A *complete* answer to the question, "Who left the party?" corresponds to one cell of this partition and thereby excludes all the other cells. An *incomplete* or *partial* answer, however, does not exclude all cells except for one, but it does exclude at least one cell. In this sense, "Either Maria or Joseph left the party" constitutes a partial answer insofar as it excludes that nobody left. Besides complete and partial answers, we can also specify over-informative answers, answers that are not complete though they imply a complete answer. Thus, the sentence, "Joseph left the party without his shoes" constitutes an over-informative answer.

---

[29] Please note that these authors spell out the specified general idea very differently.

[30] An answer is not congruent if its meaning is not an element of the meaning of *?p*. Take a look at our example: "Who left the party?" First, there are many sentences that would not even count as possible answers: e.g. "Berlin is the capital of Germany", "It's raining". To react to the question by uttering these sentences would be completely off the mark or would introduce some kind of implicature. Thus, the meanings of these sentences are obviously not elements of the meaning of the question. Second, there are other sentences that might constitute acceptable reactions to the question but still cannot be considered answers: e.g. "I do not know"; "I won't tell you". Thus, the meanings of these sentences are also not an element of the meaning of the question. Third, there are some sentences that can be considered answers, but not *congruent* answers: e.g. "Joseph left without his shoes"; "Either Joseph or Maria left". The first sentence constitutes an over-informative and the second an under-informative or incomplete answer to the question. I will come back to over-informative and incomplete answers in the main body of the text.

[31] In his seminal paper on questions, Hamblin already suggested this option (cf., Hamblin 1958).

Within such a framework, two kinds of presuppositions of questions can be differentiated, namely semantic and pragmatic presuppositions. By "semantic presupposition of a question" I mean a proposition that is implied by every complete answer-proposition (cf. Higginbotham 1996: 375). In our example, the proposition that there is a party taking place is a semantic presupposition of the question, "Who left the party?" Every complete answer-proposition implies that there is a party taking place. Thus, all worlds in which no party is taking place are excluded. This can be illustrated by the following figure (Fig.1). (Please note that square brackets are used to refer to the proposition that is expressed by the sentence inside the brackets.)
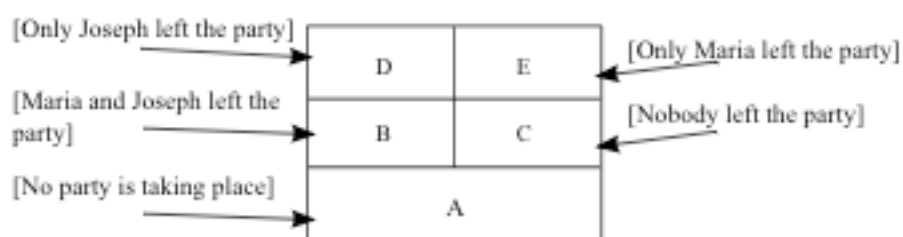

Fig. 1

The big rectangle that incorporates compartments A–E (each of which represents a set of worlds) symbolizes the space of all logically possible worlds. In a certain sense, the meaning of the question, "Who left the party?" divides this space into an answer-space and a presupposition-space. The answer-space is the exhaustive set of answer-propositions. It is divided into non-overlapping cells B–E, which together fill the whole answer-space. The presupposition space A includes all worlds that are excluded by the semantic presuppositions of the question, i.e. it includes all worlds in which no party is taking place. In more general terms, the presupposition-space can be characterized as the union of all propositions incompatible with the semantic presuppositions of the question.

Besides semantic presuppositions, we also need to introduce pragmatic presuppositions. Note that in some situations the sentence, "Maria's husband left the party", is just as good an answer as, "Joseph left the party". However, the propositions (the sets of worlds) that correspond to these two sentences are not identical. After all, there are some possible worlds in which Maria and Joseph are not married. Even though "Maria's husband left the party" and "Joseph left the party" are not equivalent, in the

sense that these sentences do not express the same proposition, the sentences can nevertheless in certain circumstances constitute equally good answers to the given question. How is that possible?

Building on insights by Robert Stalnaker (1970, 1973), this can be explained by recourse to beliefs taken for granted by the participants of the conversation. Assume that all participants of a conversation take for granted that Maria and Joseph are married. Thus, all possible worlds in which they are not married are excluded in the sense that they are not an element of the sets of possible worlds that represent the doxastic sets of the participants. In this case the conversational impact of the answers, "Joseph left the party", and, "Maria's husband left the party", will be identical. This can be illustrated by the following figure (Fig. 2).
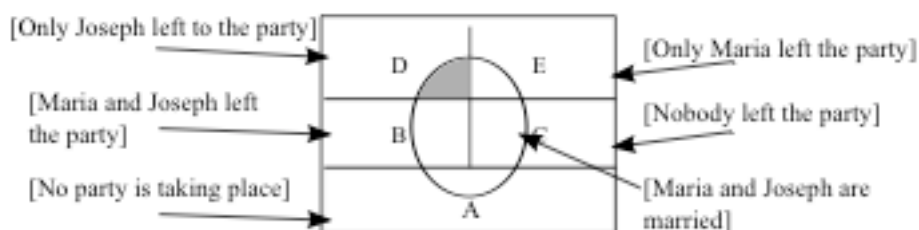

Fig. 2

The circle in the middle symbolizes the set of worlds in which Maria and Joseph are married, and thereby it symbolizes a belief taken for granted by the participants—i.e. all worlds in which they are not married (all worlds outside the circle) are excluded. Fig. 2 represents that the set of worlds in which Maria and Joseph are married contains elements of the sets A, B, C, D, and E. In some worlds in which Maria and Joseph are married, no party is taking place; in others a party is taking place and both left the party; still in others, nobody left, etc. If the participants of the conversation take for granted that Maria and Joseph are married, then the answers, "Joseph left the party", and, "Maria's husband left the party", will both correspond to the same set of worlds, namely, the set of worlds represented in Fig. 2 by the grey-tinted compartment representing the intersection of the propositions [Only Joseph left the party] and [Maria and Joseph are married]. Thus, by introducing pragmatic presuppositions and characterizing them as the set of beliefs taken for granted by the participants it can be explained why the answers, "Joseph left the party", and, "Maria's husband left the

party", can in certain circumstances constitute equally complete and satisfying answers.

Gronendijk and Stokhof, therefore, relativize the meaning of questions on the set of beliefs taken for granted by the participants of a conversation (Gronendijk & Stokhoff 1984: 147–148). Let us call the union of the sets of possible worlds corresponding to the beliefs *all* participants take for granted the "context set". The meaning of an interrogative sentence *?p* can then be specified as the partition of the *context set* into an *exhaustive set of answer-propositions* on the one hand and the union of all *propositions incompatible with the semantic presuppositions* of *?p* on the other.

This characterization of questions and their semantic and pragmatic presuppositions will help us to specify the structure and presuppositions of cognitive projects, which in turn will allow us to spell out the idea that beliefs regarding the presuppositions of cognitive projects are promising candidates of non-truth-indicatively justified beliefs.

4.2. The characterization of cognitive projects

Based on the proposition-set theory, we can spell out the idea that cognitive projects are closely related to questions in more detail.

What is a cognitive project? A cognitive project is an attempt to formulate and answer a question. The meaning of a question in the context of a cognitive project can be understood as a partition of the set of possible worlds along the lines of the proposition-set theory.[32] The meaning of the question gives the cognitive project its direction and goal. The goal of a cognitive project consists in identifying the complete and true answer to the question, thereby excluding all other possible answers.[33] [34]

---

[32] Please note that this does not presuppose that the proposition-set theory is correct as a semantic theory of interrogative sentences. Even if one thinks that the proposition-set theory is unconvincing as a semantic theory of everyday interrogative sentences, the theory could nevertheless be used as a tool to analyze the questions in the background of cognitive projects.

[33] This specification of the goal of a cognitive project is given from a veritistic perspective. After all, the given characterization of cognitive projects is supposed to lead to an interesting version of *veritistic* epistemic consequentialism that is not committed to (TI). All philosophers claiming that truth (or more precisely improving the balance of true over false beliefs) is not the only and not even the most important epistemic value will probably dismiss this specification of the goal of cognitive projects.

[34] Note that two cognitive projects described by interrogative sentences with the same wording might still be very different projects. This is because two interrogative sentences with the same wording can have a very different meaning, i.e. they can partition the set of possible worlds very differently.

What does it mean to make progress within a cognitive project? One has made progress in a cognitive project if one (or more) of the following conditions are met: (i) the set of possible answers is identified (or at least *some* possible answers are identified); (ii) a true but incomplete answer is identified, i.e. at least one possible answer is excluded; or (iii) the true and complete answer is identified, i.e. all other possible answers are excluded.

Meeting condition (i) in a certain sense constitutes minimal progress within a project because by meeting (i) only the meaning of the question (i.e. the set of possible congruent answers) that constitutes the project in the first place is clarified. Meeting condition (iii) constitutes maximal progress within a cognitive project because by meeting (iii) the project is successfully concluded.

The given characterization of a cognitive project makes it reasonable to regard such a project as a systematic attempt to achieve the epistemic aim of improving the balance of *interesting* true beliefs over false ones. Particularly in light of the double character of this aim, performing a cognitive project seems to be one of the most attractive *means available* to achieve this aim. In light of the double character of the epistemic goal, we are supposed to strive for increasing the number of interesting true beliefs (beliefs in the truth of which we are interested), but we have to strive for them in a way that simultaneously avoids false beliefs. This is exactly what we do during the performance of a cognitive project. First, we ask for something of interest to us. Second, we try to clarify the meaning of the question by clarifying what would count as a possible answer to it. Third, we systematically check each possible answer and dismiss it if it is false. In the best case, we will conclude such a project by having identified the one and only true and complete answer to the question we are interested in.

What are presuppositions of a cognitive project? The presuppositions of a cognitive project are the semantic and pragmatic presuppositions of the question constituting the project. Semantic presuppositions are the propositions implied by all complete answers to the question constituting the project. Pragmatic presuppositions are the beliefs taken for granted by all people engaged in the project, or more precisely: the propositions the truth of which is taken for granted by the participants of a project.

The latter presuppositions have two important functions. First, remember that the meaning of a question is specified as the partition of the context set. Thus, the propositions the truth of which is taken for granted by the participants of a cognitive

project in part determine the meaning of the question that defines the structure and direction of the project. Second, pragmatic presuppositions of a cognitive project will specify in which circumstances a possible answer qualifies as being excluded. It will thereby in part determine in what circumstances progress within a project is achieved. Suppose Sherlock Holmes and Dr. Watson are engaged in a cognitive project specified by the question, "Who committed the murder?" They systematically check the possibly correct answers: "the gardener committed the murder", "the chauffeur committed the murder", etc. Let us suppose that both believe that, if a person has been a thousand miles away from the crime scene at the time of the murder, then that person did not commit the murder. The content of this conditional belief is part of the context set i.e. it is a pragmatic presupposition of the question that defines the project. It is obvious that this pragmatic presupposition will determine under which circumstances Holmes and Watson are in a position to exclude a possible answer, thereby making progress in their project.

When is a cognitive project hopeless or irrational? If the presuppositions of a cognitive project imply that no progress within a project can be made or if the presuppositions are contradictory, then the cognitive project is hopeless or irrational.

4.3. Cognitive projects and non-truth-indicatively justified beliefs

Based on the characterization of cognitive projects given in subsection 4.2, we can spell out a version of veritistic epistemic consequentialism that allows for non-truth-indicatively justified beliefs and is therefore not committed to (TI) in its general form. At the beginning of section 4, we saw that veritistic consequentialists who want to allow for non-truth-indicatively justified beliefs have to give satisfying answers to questions (1)–(3).

Question (1): Which beliefs have a chance of being non-truth-indicatively justified? The beliefs in presuppositions of cognitive projects that are neither hopeless nor irrational are candidates for being non-truth-indicatively justified beliefs; i.e. a belief in presuppositions of such a project has a chance of being justified independently from truth-indicative factors speaking in favor of the truth of the belief itself. (Thus, these beliefs constitute possible exceptions to (TI)). Please note that the candidates for

non-truth-indicatively justified beliefs are restricted to beliefs in presuppositions of projects that are neither *hopeless* nor *irrational;* the presuppositions of any old cognitive project will not do. Note also that beliefs in presuppositions of such projects are only *candidates* of non-truth-indicatively justified beliefs. Actually to count as non-truth-indicatively justified, these beliefs have to satisfy further conditions, as explained below.

Question (2): In which sense are non-truth-indicatively justified beliefs epistemically justified, i.e. in which sense are they conducive to our epistemic aim of improving the balance of (interesting) true beliefs over false ones? The given characterization of cognitive projects has revealed that a cognitive project can be considered a systematic quest to achieve the epistemic goal of increasing the number of intersting true beliefs while at the same time avoiding false ones. In fact, engaging in cognitive projects seems to be one of the best means available to achieve this goal. Because beliefs in the presuppositions of cognitive projects are necessary to perform the projects—as the presuppositions constitute the structure and direction of these projects in the first place—and performing these projects in turn seem to be one of the best systematic means available to achieve the specified epistemic goal, beliefs in the presuppositions of cognitive projects are conducive to our epistemic goal. [35]

However, please note that this is only true from a first-person perspective. It is not guaranteed that performing a cognitive project, which was made possible by holding certain presuppositions to be true, actually leads to a greater balance of interesting true beliefs over false ones. However, subjects engaged in such projects expect and intend that this is the case. Indeed, one of the main reasons subjects are engaged in cognitive projects is that by performing these projects they intend to dismiss false answers and eventually find the complete true answer to a question in which they are interested. Thus, the envisaged veritistic consequentialist position that does not subscribe to the truth-indicative principle (TI) is a variant of *subjective* consequentialism insofar as not the actual consequences of holding a non-truth-indicatively justified belief, but rather the consequences intended and expected by the subject are relevant.

Question (3): Is there a way to spell out the conduciveness of non-truth-

---

[35] This answer can also be considered an answer to an important challenge raised with respect to Wright's theory of entitlement, namely the challenge to specify in which sense "entitlement" is is an epistemic notion rather than a purely pragmatic one (cf. Pritchard 2007: 207; Jenkins 2007: 27). Wright's own answer to this challenge proceeds broadly along similar lines (cf. Wright 2014: 238-239).

indicatively justified beliefs towards the specified epistemic goal that respects constraints (a) and (b)? The conduciveness of non-truth-indicatively justified beliefs towards the epistemic goal is spelled out in a way that meets the means-end relation (see constraint (a)). Believing the presuppositions of a cognitive project is a necessary means to perform such a cognitive project, and performing such a project, in turn, seems to be one of the best means available to achieve the specified epistemic goal. Thus, the means-end relation between holding a non-truth-indicatively justified belief in the presuppositions of cognitive projects and the specified epistemic aim is neither arbitrary nor remote.

What about the no-trade-off constraint (see constraint (b))? This constraint demands that it should not be allowed that we knowingly make epistemic sacrifices with respect to a belief in *p* to achieve a greater epistemic good with respect to beliefs in other propositions. The truth-fairy example vividly illustrates the importance of this constraint. By subscribing to (TI), epistemic consequentialists can ensure that this constraint is met. However, it is unclear whether the envisaged consequentialist position that rejects (TI) and allows that beliefs in presupposition of cognitive projects are non-truth-indicatively justified satisfies the constraint as well. The easiest way to make sure that the envisaged position satisfies constraint (b) consists in adding further conditions for non-truth-indicatively justified beliefs. Thus, the envisaged account should be modified and extended along the following lines:

> *S*'s belief that *p* is non-truth-indicatively justified if and only if
>
> (I) *p* is a presupposition of one of *S*'s cognitive projects that is neither hopeless nor irrational;
>
> (II) the body of evidence *E* available to *S* does not make it likely that *not-p*;
>
> ⋮

Condition (I) ensures that non-truth-indicatively justified beliefs are conducive to the epistemic aim in a way that satisfies the means-ends relation (a). Condition (II) ensures that the no-trade-off constraint (b) is met as well. Condition (II) ensures that one cannot gain a non-truth-indicatively justified belief by knowingly making epistemic sacrifices with respect to a belief in a certain proposition to achieve a greater epistemic good with

respect to beliefs in other propositions. Thus, even though the proposed variant of veritistic epistemic consequentialism is not committed to (TI), it nevertheless satisfies constraints (a) and (b).

The dots in the definition indicate that the account could easily be extended by further conditions (III), (IV), etc. Conditions (I) and (II) already seem to suffice to circumvent all counter-examples against epistemic consequentialism put forward in the literature. However, conditions (I) and (II) by themselves still seem too permissive, insofar as they probably make it too easy to have a non-truth-indicatively justified belief. In order to respond to this problem, it might become necessary to extend the definition by further conditions.[36] [37]

As far as the definition in its current form goes, a promising example of a non-truth-indicatively justified belief is the belief that my senses work reliably. The proposition that my senses are working reliably is unquestionably a presupposition of many of my cognitive projects that are neither hopeless nor irrational, and the body of evidence available to me does not make it likely that this proposition is false.[38]

In summary, the specified version of epistemic consequentialism has the following features. First, it is *veritistic* insofar as it claims that beliefs can be epistemically evaluated by recourse to the specified truth-goal. More precisely, it holds that, if a belief is epistemically justified, then holding the belief is conducive to the aim of increasing the number of interesting true beliefs and simultaneously avoiding false beliefs. Second, the position is *non-truth-indicative* insofar as it allows that at least *some* beliefs are justified independently of truth-indicative factors' speaking in favor of the truth of the beliefs themselves. Third, the position is *subjective* insofar as—at least with respect to non-truth-indicatively justified beliefs—not the actual consequences of holding the belief are epistemically relevant, but rather the consequences the subject

---

[36] Of course, the definition of "non-truth-indicative justification" cannot be extended by conditions demanding truth-indicative factors (i.e. evidence for the belief or the reliability of the belief-forming process). Otherwise, it would not be a definition of *non-truth-indicative* justification.

[37] Building on Wright's theory of epistemic entitlement, we could, for example, add the following condition: (III) any attempt to justify *p* by evidence would be epistemically circular insofar as it would involve something of no more secure standing than *p* itself (cf. Wright 2004: 191–192).

[38] Please note that, even if we add condition (III) to the definition of non-truth-indicative justification (see fn.36), the belief that my senses work reliably would still seem like a promising example of a non-truth-indicatively justified belief. It seems plausible that my belief that my senses work reliably cannot be justified *a priori*. However, any attempt to justify the belief by empirical evidence seems to be epistemically circular insofar as it would presuppose that my senses work reliably. Therefore, the belief in question seems to satisfy condition (III) as well.

intends or reasonably expects.[39] Engaging in a cognitive project does not guarentee that the balance of true over false beliefs is improved. However, epistemic agents expect that this is the case, the only reason why they engage in these kind of projects is because they intend to make headway with respect to the aim of increasing the number of interesting true beliefs while at the same time avoiding false ones. Fourth, even though the position rejects (TI) as a general principle, it nevertheless satisfies the means-end constraint (see constraint (a)) and the no-trade-off constraint (see constraint (b)).

5. Concluding Remarks

The results of the discussion can be summarized as follows: Veritistic epistemic consequentialism is structurally analogous to hedonistic consequentialism in ethics. Both variants of consequentialism explicate normative notions via recours to final values and ultimate goals in their respective domains. Furthermore, consequentialism in both domains can be spelled out in a *direct* or *indirect* and a *subjective* or *objective* version (see. Sect. 2).

Besides these similarities, there are also important dissimilarities. One difference consists in the fact that the former has to meet the means-end constraint (see constraint (a)) and the no-trade-off constraint (see constraint (b)), whereas the latter does not have to satisfy analogous constraints in the ethical domain. Epistemic consequentialists can ensure that their theory meets the specified constraints by incorporating the truth-indication principle (TI) into their consequentialist framework, i.e., by claiming that a belief is epistemically justified only if there are factors indicating the truth of the belief *itself*. This is why popular versions of epistemic consequentialism subscribe to (TI), whereas popular versions of ethical consequentialism do not subscribe to an analogous pleasure-indication principle (PI)* (see Sect. 3).

However, building on ideas suggested by Wittgenstein, Williams, and most explicitly by Wright, a version of veritistic epistemic consequentialism can be defended that rejects (TI) as a general principle, but is nevertheless able to satisfy constraints (a) and (b). The basic thought is roughly that beliefs in presuppositions of cognitive

---

[39] Whether the same is true for truth-indicatively justified beliefs as well is left open by the suggested version of non-truth-indicative epistemic consequentialism.

projects can be justified independently of factors speaking in favor of their truth. Nonetheless, these beliefs can be *epistemically* justified in the sense that they are conducive to our epistemic aim of improving the overall balance of (interesting) true beliefs over false ones—at least from a first-person perspective. Combining the idea that a cognitive project is an attempt to answer a question with the proposition-set theory allows one to specify this basic thought in more detail. Furthermore, this combination delivers the resources to spell out the conduciveness of non-truth-indicatively justified beliefs regarding the specified epistemic aim in a way that respects constraints (a) and (b) (see Sect. 4).

Thus, even though the truth-indication principle (TI) seems plausible for many beliefs, veritistic consequentialists can allow for exceptions to the principle without thereby violating contstraints (a) and (b). Hence, just as hedonistic consequentialism in ethics is not committed to the pleasure-indication principle (PI)*, so veritistic consequentialism in epistemology is not committed to the truth-indication principle (TI).[40]

References

Alston, William P. (1989). *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press.

Alston, William P. (2005). *Beyond "Justification": Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.

Berker, Selim (2013a). "Epistemic Teleology and the Separateness of Propositions." *Philosophical Review*, 122: 337–392.

Berker, Selim (2013b). "The Rejection of Epistemic Consequentialism." *Philosophical Issues*, 23: 363–387.

BonJour, Laurence (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

BonJour, Laurence (2003). "A Version of Internalist Foundationalism." In Bounjour and Sosa: *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*. Malden, MA: Blackwell.

Conee, Earl (1992). "The Truth Connection." *Philosophy and Phenomenological Research*, 52: 657–669.

Dewey, John (1938). *Logic—The Theory of Inquiry*. New York: Holt, Rinehart, and Winston.

Feldman, Richard (2000). "The Ethics of Belief*." Philosophy and Phenomenological Research*, 92: 1–24.

Foley, Richard (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.

Foley, Richard (1993). *Working without a Net: A Study in Egocentric Epistemology*. Oxford: Oxford University Press.

Fumerton, Richard (2001). "Epistemic Justification and Normativity." In Steup (ed.): *Knowledge, Truth, and Virtue*. Oxford: Oxford University Press.

Goldman, Alvin (1979). "What is Justified Belief?" In Pappas (ed.): *Justification and Knowledge*. Dordrecht: Reidel.

Goldman, Alvin (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

Goldman, Alvin (2001). "The Unity of the Epistemic Virtues" In Fairweather and Zagzebski (eds): *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford University Press.

Grimm, Stephen (2008). "Epistemic Goals and Epistemic Values." *Philosophy and Phenomenological Research*, 77: 725–744

Grimm, Stephen (2009). "Epistemic Normativity." In Haddock, Millar, and Pritchard (eds.): *Epistemic Value*. Oxford: Oxford University Press.

Groenendijk, Jeroen and Stokhof, Martin (1982). "Semantic Analysis of wh-complements." *Linguistics and Philosophy* 5: 175-233.

Groenendijk, Jeroen and Stokhof, Martin (1984). "On the Semantics of Questions and the Pragmatic of Answers". In Landman and Veltman (eds.): *Varieties of Formal Semantics*. Dordrecht: Foris.

Groenendijk, Jeroen and Stokhof, Martin (1997). "Questions." In van Bentham, and ter Meulen (eds.): *Handbook of Logic and Language*. Amsterdam: Elsevier.

Haddock, Adrian; Millar, Alan; and Pritchard, Duncan (ed.) (2009): *Epistemic Value*. Oxford: Oxford University Press.

Hamblin, C. L. (1958). "Questions." *Australasian Journal of Philosophy*, 36: 159–168.

Hamblin, C. L. (1973). "Questions in Montague Grammar." *Foundations of Language*, 10: 41–53.

Higginbotham, James (1996). "The Semantics of Questions." In Lappin (ed.): *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell.

Hookway, Christopher (1996). "Questions of Context." *Proceedings of the Aristotelian Society*, 96: 1–16.

James, William (1897). "The Will to Believe." In *The Will to Believe and Other Essays in Popular Philosophy*. New York: Dover.

Jenkins, Carrie S. (2007). "Entitlement and Rationality." *Synthese*, 157: 25–45.

Karttunen, Lauri (1977). "Syntax and Semantics of Questions." *Linguistics and Philosophy*, 1: 3–44.

Kvanvig, Jonathan (2005). "Truth is not the Primary Epistemic Goal." In Steup, and Sosa (eds.): *Contemporary Debates in Epistemology*. Malden, MA; Blackwell.

Lehrer, Keith (1974). *Knowledge.* Oxford: Oxford University Press.

Leitgeb, Hannes and Pettigrew, Richard (2010): An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science*, 77: 201–235

Littlejohn, Clayton (2012). *Justification and the Truth-Connection*. Cambridge: Cambridge University Press.

Pritchard, Duncan (2007). "Wittgenstein's *On Certainty* and Contemporary Anti-Scepticism." In Moyal-Sharrock and Brenner (eds.): *Readings of Wittgenstein's On Certainty.* Basingstoke: Palgrave Macmillan.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Sinnott-Armstrong, Walter (2014). "Consequentialism", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/consequentialism/>.

Stalnaker, Robert (1970). "Pragmatics." *Synthese*, 22: 272–289.

Stalnaker, Robert (1973). "Presuppositions." *Journal of Philosophical Logic*, 2: 447–457.

Thomson, Judith Jarvis (1976). "Killing, Letting Die, and the Trolley Problem." *The Monist*, 59: 204–217.

Williams, Michael (1996). *Unnatural Doubts*. Princeton: Princeton University Press.

Williams, Michael (2001). *Problems of Knowledge*. Oxford: Oxford University Press.

Wittgenstein, Ludwig (1996). *On Certainty*. Anscombe and von Wright (eds.), Anscombe and D. Paul (trans.), Oxford: Blackwell.

Wright, Crispin (2004). "On Epistemic Entitlement: Warrant for Nothing (and Foundations for Free)." *Aristotelian Society Supplementary*, 78: 167–212.

Wright, Crispin (2014). "On Epistemic Entitlement (II) — Welfare State Epistemology." In Dodd and Zardini (eds.): *Scepticism and Perceptual Justification*. Oxford: Oxford University Press.