KİLİKYA FELSEFE DERGİSİ Sayı: 2, Ekim 2018, ss. 17-34
CILICIA JOURNAL OF PHILOSOPHY Issue: 2, October 2018, pp. 17-34

Makale Geliş Tarihi | **Received:** 08.04.2017 E-ISSN: 2148-9327
Makale Kabul Tarihi | **Accepted:** 07.09.2018 http://philosophy.mersin.edu.tr
Araştırma Makalesi | **Research Article**

IS THE TRUTH CONDITION SUPERFLUOUS FOR DEFEASIBILITY THEORIES OF KNOWLEDGE?

Uğur AYTAÇ*

Abstract: Defeasibility theories aim to reach a plausible definition of knowledge by finding strategies to exclude true beliefs based on faulty justifications. Different philosophers have advanced with their own understandings of undefeated justification. Zagzebski (1994) indicates that the strong defeasibility condition violates independence between truth and justification because undefeated justification never leads to false beliefs. Following this, Zagzebski and some other philosophers who pursue a similar line of reasoning (e.g., Merricks, 1995) conclude that undefeated justification entails truth. In this paper, I argue that the truth condition is not superfluous by presenting an example of undefeated justification that does not entail truth. My claim is that beliefs about metaphysical questions (e.g., Does God exist?) can have undefeated justifications. Nonetheless, such undefeated justifications are not capable of assigning truth to the beliefs that they support.

Keywords: Analysis of Knowledge, Defeasibility, Justification, Epistemology, Truth

DOĞRULUK KOŞULU BİLGİNİN SARSILABİLİRLİK KURAMLARI İÇİN GEREKSİZ MİDİR?

Öz: Sarsılabilirlik kuramları makul bir bilgi tanımına ulaşabilmek için hatalı gerekçelendirmelere dayalı doğru inançları eleme stratejisini güder. Farklı felsefeciler kendilerine özgü sarsılabilirlik anlayışları geliştirmiştir. Sarsılmaz gerekçelendirmeler hiçbir zaman yanlış inançların dayanağı olmadığından Zagzebski (1994) katı sarsılabilirlik şartının inancın doğruluğu ile gerekçelendirilmesi arasındaki bağımsızlığı ortadan kaldırdığını ifade eder. Bunu takiben, Zagzebski ve yine benzer görüşleri savunan bazı felsefeciler (örneğin, Merricks, 1995) sarsılmaz gerekçelendirmenin zorunlu olarak doğruluğu beraberinde getirdiği sonucuna varır. Bu makalede, sarsılmaz gerekçelendirmelerin her zaman doğru inanca ulaşmaması durumuna örnek göstererek, doğruluk koşulunun lüzumsuz olmadığını iddia ediyorum. İddiama göre

* Doktora Öğrencisi | Ph.D. Student
Amsterdam Üniversitesi, Amsterdam Sosyal Bilimler Araştırma Enstitüsü, Siyaset Bilimi, Hollanda | University of Amsterdam, Amsterdam Institute for Social Science Research, Political Science, The Netherlands
u.aytac@uva.nl
Orcid Id: 0000-0002-4773-5472

Aytaç, U. (2018). Is the Truth Condition Superfluous for Defeasibility Theories of Knowledge? ***Kilikya Felsefe Dergisi*****, (2), 17-34.**

metafizik sorulara (örneğin: Tanrı var mıdır?) ilişkin inançlar sarsılmaz gerekçelendirmelere sahip olabilirler. Buna rağmen söz konusu gerekçelendirmeler destekledikleri inançlara doğruluk atfedemez.

Anahtar Kelimeler: Bilgi Kavramının Analizi, Sarsılabilirlik, Gerekçelendirme, Bilgi Kuramı, Doğruluk

1. Introduction

Defeasibility theories are one set of attempts to analyze knowledge while tackling the frustration posed by the famous Gettier (1963) cases. These theories modify the classical definition of knowledge as *justified true belief* (JTB) by requiring that justification is indefeasible. Indefeasibility is a condition such that the justification cannot possibly be undermined. By this means, defeasibility theories exclude the counter-examples to the JTB account through removing falsely justified true beliefs from their conception of knowledge. Different philosophers have developed their own understandings of undefeated justification. For instance, Zagzebski (1994) indicates that the strong defeasibility condition violates independence between truth and justification because undefeated justification never leads to false beliefs. Following this, Zagzebski and some other philosophers who pursue a similar line of reasoning (e.g., Merricks, 1995) conclude that undefeated justification entails truth. One of the most important implications of this view is that the truth condition is superfluous in defeasibility theories since the truth of a belief is necessarily guaranteed by the infallible nature of justification. In this paper, I argue that the truth condition is not superfluous by presenting an example of undefeated justification that does not entail truth. My claim is that beliefs about metaphysical questions (e.g., Does God exist?) can have undefeated justifications although these justifications are not capable of assigning truth to the beliefs in question. What I mean by undefeated justification of metaphysical beliefs relies on a slight modification of Klein's (1980) defeasibility condition. According to this modification, the truth of defeaters must be accessible to the epistemic agents with normal rational capabilities under optimal conditions. I call this *the accessibility condition*. Following this, my claim is that some justifications of metaphysical beliefs are able to remain undefeated since no accessible defeater genuinely undermines them.

In the first section, I review Klein's defeasibility theory of knowledge. Then, in the second section, I present Zagzebski's views on the violation of independence between truth and justification. In the third section, I introduce the accessibility condition and argue that it is coherent with the logic of defeasibility theories of knowledge. In the fourth section, I contend that some metaphysical propositions may have undefeated justifications even though we do not know these propositions. I refer to the debate between atheists and theists on the cosmological argument regarding the origin of the universe as an example of this sort of belief. It is completely possible that both theories (atheism and theism) include no false empirical beliefs or logical fallacies in their justifications. In such a case, both of them still remain justified since the controversy between atheistic and theistic beliefs relies on divergences of our intuitions about

specific notions that are used in the justifications. However, these conflicts of intuitions do not render genuine defeaters possible since the accessibility condition is not met.

Before starting the main discussion, let me make a final remark about the fittingness of my argument for the subject matter at hand. Certain defeasibility theorists explicitly state that their account of knowledge only aims to explain empirical knowledge (Klein, 1984, p. 14). Hence, one might claim that my argument is not a proper treatment of defeasibility theories since it is based on *a priori* justifications of metaphysical beliefs. In contrast, some epistemologists hold that "epistemologists seek to understand all actual or possible knowledge, not just some of it" (Hetherington, n.d.). Yet, I do not wish to engage in a debate about what specific task defeasibility theories ought to adopt. It is entirely understandable why Klein only focused on empirical knowledge in his book given that he aimed to refute the skeptical challenge regarding empirical beliefs. However, his intention to narrow the scope of his theory does not mean that defeasibility theories are incompatible with *a priori* justification. His preference is just a product of what he conceives as *relevant* in relation to the skeptical attack on empirical knowledge. The intellectual context of my paper is different. I try to apply Klein's definition of defeaters to a much broader context. By this means, it would be possible to have an account of defeasibility as a general theory of knowledge, containing both *a priori* and *a posteriori* knowledge. In this sense, one might think of my argument as a preliminary attempt to extend the limits of defeasibility theories and derive a general theory of knowledge from its conventional forms.

2. Defeasibility Theories of Knowledge

Defeasibility theories aim to reach a plausible definition of knowledge by finding strategies to exclude true beliefs based on faulty justifications (Klein, 1980; 1984; Lehrer & Paxson, 1969; Levy, 1977). It is usually asserted that S knows p if and only if (i) p is true, (ii) S believes that p, (iii) S is justified in believing that p, (iv) there is no other statement that defeats this justification (Lehrer & Paxson, 1969, p. 227). I focus on one of the prominent examples of these theories: Klein's account of defeasibility. Klein proposes the following criterion for the defective justification: "the justification of p by evidence e is defective if and only if there is some true proposition, d, such that (d & e) fails to justify p" (Klein, 1984, p. 125).

Consider the famous Tom Grabit example. An epistemic agent, S, sees a person, who looks like Tom Grabit, removing a book from the library by concealing it under his coat. Based on this evidence, S believes that Tom Grabit stole the book. Assume that S's evidence can be considered reliable. However, unbeknownst to S, Tom has an identical twin, John, who was in the library on the day at issue. Even if it is true that Tom stole the book, S's available evidence does not seem to justify S's belief in light of the true statement that Tom has an identical twin who was in the library on that day. If S knew that Tom has an identical twin, she would need some additional evidence to believe that Tom, rather than John, stole the book. This example can be reformulated according to Klein's criterion for defective justification:

p: Tom stole the book.

e: Tom-like person removed the book by concealing it under his coat.

d: Tom has an identical twin, who was in the library on the day at issue.

The conjunction (e & d) fails to justify p because d implies that e is not as reliable as initially believed. Since the evidence relies on the alleged match between Tom Grabit and the thief by virtue of their similar appearance, a true statement affirming that there is another person, who looks exactly alike undermines the justification. True statements that undermine available evidence are called defeaters (Sudduth, n.d.). In the previous example, d is a defeater because it undermines the available evidence.

Klein (1980) makes an important distinction between misleading and genuine defeaters. Misleading defeaters derive their capacity to undermine a justification from a false proposition (Klein, 1980, pp. 84-85). For instance, consider the second version of the Tom Grabit example. Suppose that d is a false statement. Tom does not have an identical twin. We have another defeater, d*: Tom's mother says that d. Assume that S believes that Tom's mother's testimony is a reliable source of evidence. Therefore, d* renders d plausible. The conjunction (e & d) fails to justify p. Moreover, d* is true since Tom's mother really said that d. However, suppose that, unbeknownst to S, Tom's mother is a pathological liar. Her saying that d is just one of her lies. In this case, d* is a misleading defeater since it meets the following conditions (Klein, 1980, p. 85):

(i) d* is true,

(ii) d* renders plausible a false proposition d,

(iii) d is such that (e & d) fails to justify p,

(iv) d* defeats the justification only because it renders d plausible.

Klein holds that a justification, which is undermined by a misleading defeater, is restorable because there is another true statement that cancels out the impact of the misleading defeater (Klein, 1980, p. 86). What he means by restoration is a sort of reactivation of the justificatory power of the initial evidence. Let me illustrate this point below:

p, e and d are the same.

d*: Tom's mother says that d.

r: Tom's mother is a serial liar.

Even though (e & d*) fails to justify p, the initial justification based on e is restorable. Consider that S realizes that r. Since r implies that Tom's mother's testimony is not a reliable source of evidence, d* loses its power to undermine the initial justification, given that d* derives that power from S's trust in Tom's mother's testimony. In other words, r cancels out the impact of d*. Consequently, e suffices to justify p. In this sense, the initial justification is restored. In contrast, there is no other true statement that cancels out the

impact of genuine defeaters. In the first version of the Tom Grabit example, Tom really has an identical twin. To justify p, there must be some additional evidence. For instance, if S finds out that Tom's brother has never been in the room where S saw the Tom-like person, then S would be justified to believe that p. However, this justification does not solely rely on the initial evidence:

> p, e and d are the same.
>
> e*: Tom's brother, John, has never been in the room where S saw the Tom-like person.

d is a genuine defeater because (e & d) fails to justify p whereas no true statement can resolve this failure. The same conclusion, p, can be held only if there is an additional piece of evidence, e*. Klein contends that "S's true justified belief is knowledge if and only if all defeaters of S's justification are misleading defeater" (Klein, 1980, p. 84). A justification is not *really* defeated if its defeat relies on a false proposition.

3. The Claim That Undefeated Justification Entails Truth

Zagzebski (1994) makes a distinction between strong and weak defeasibility conditions in defeasibility theories. The strong defeasibility condition implies that a justification is undermined if there is any true proposition that counts as evidence against that justification. Klein's account of defeasibility can be classified under this category. According to Klein, any true statement, which does not render a false proposition plausible, is a genuine defeater when its conjunction with available evidence fails to justify p. His account parallels what Zagzebski calls the strong defeasibility condition. Zagzebski holds that it is not possible for a false belief to be justified according to the strong defeasibility condition (Zagzebski, 1994, p. 70). A false belief cannot be supported by a set of evidence because either certain elements of the empirical evidence are false or (if it is a deductive argument) there is a logical fallacy in the sense that the conclusion, p, does not follow from the evidence.[1] Otherwise, it would not be possible to justify a false belief. In both cases, there is a true proposition informing us about the falsehood of empirical evidence or the logical fallacy. As a result, any justification of a false belief will be undermined by these true statements. Hence, it is asserted that undefeated justification always leads to true belief.

The claim that undefeated justification entails truth is relevant to the analysis of knowledge in a specific way. If undefeated justification entails truth, there must be an explanation for why defeasibility theorists still need to use the truth condition in their analysis of knowledge. Lehrer, a defeasibility theorist, already explicitly states that the truth condition is trivial (1990, p. 171). However, another defeasibility theorist, Klein (1984), refers to a distinct truth condition in his work. Yet, it seems that Lehrer has a point. Once you accept the strong defeasibility condition, it *prima facie* seems that knowledge can be defined as *undefeated justified belief* without a distinct truth condition.

[1] Keep in mind that I do not discuss other forms of justification such as abduction to avoid further complexities in my argument. I mainly consider deductive and inductive reasoning.

Since it is ostensibly impossible for an undefeated justification to lead to a false belief, one can argue that using the truth condition in the analysis of knowledge is superfluous. An undefeated justification allegedly leads to a true belief anyway. In the rest of the paper, I am going to contend that the truth condition is not superfluous. To ground my claim, I will first further elaborate on the nature of defeaters.

4. Undefeated Justification Revisited: The Accessibility Condition

In order to illustrate defeaters, epistemologists construct numerous thought experiments. In this section, my aim is to reinterpret this way of illustrating defeaters as a part of the nature of defeaters. In other words, I claim that a property of defeaters is inseparable from our thinking about them through our thought experiments. These thought experiments rely on our imagination to picture a perspective from which we have access to a true statement undermining the initial justification. I will formulate this relationship between defeaters and epistemic agents, who construct thought experiments, by *the accessibility condition* (AC):

> The AC: The truth of defeaters is accessible to ideal epistemic agents (epistemic agents with normal rational capabilities under optimal conditions).

By having access to the truth of a proposition, I mean that there is a desirable epistemic relationship between an ideal epistemic agent and a proposition. It ought to be the case that ideal epistemic agents have a sufficiently broad perspective to be able to first imagine d, and then potentially know that it is true. Klein very briefly acknowledged the possibility of creating an additional requirement that defeaters ought to be humanly knowable (1984, p. 171). In this section, I aim to further elaborate what a fuller expression of such condition would look like. Consider the early version of the Tom Grabit example. The defective nature of the justification relies on our hypothetical position as observers. First, the truth of d (Tom has an identical twin) is very easily imaginable from our point of view. There is nothing that challenges our faculty of imagination in assuming that there is such a true proposition. Moreover, we assume ourselves to be in a position of evidential superiority. We, as observers, have normal rational capabilities just like the epistemic agent, S, in the thought experiment. However, it is assumed that we observe the situation under optimal conditions unlike S. In addition to this, we assume that d is a knowable statement under optimal conditions. In other words, knowing that d is true is not beyond the reach of ideal epistemic agents. Defeaters are supposed to be accessible in this manner since the thought experiments in question cannot be relevant to our epistemic paradigm otherwise. What would be the point of criticizing S' judgment if it is not relevant to the standards of our epistemic paradigm in the first place? This critique cannot even be asserted in the absence of imageability or knowability of defeaters since one cannot reasonably expect an unimaginable or unknowable true statement to be the source of proper epistemic critique.

One might ask whether optimal conditions are the same as accessibility (i.e., having a sufficiently broad perspective from which an observation is conducive to epistemically

competent evaluation of available evidence). The difference is that the optimal conditions do not necessarily belong to the meta-level of epistemologists. It is possible to create optimal conditions within the thought experiment. Suppose that, S's friend, K, believes that Tom has an identical twin since she saw two Tom-looking persons in Tom's house when she went to Tom's birthday party. Moreover, she heard that both Tom and the other Tom-looking person call the same woman "mom". If K is added to the thought experiment, it is obvious that K is under optimal conditions in comparison to S. Following from this, one can contend that optimal conditions refer to the relative positions of advantage of epistemic agents within a single epistemic paradigm.

In contrast, I use the term 'accessibility' more to speculate on the meta-epistemic level from which philosophers (or ideal epistemic agents) look at the world. Let me give an example to clarify what I mean. Imagine that there is an alternative universe, where all rational epistemic agents lack depth perception (i.e., perceiving the external world two-dimensionally). This deficiency is an essential part of normal rational epistemic agents in this universe. Moreover, consider that all their scientific endeavors evolved in a way that they also fail to indirectly acquire the knowledge of depth (at least not in the same way as we have). Epistemologists of this universe construct a thought experiment in which S believes that p on the grounds that e supports p. Moreover, epistemologists make an assumption that there is no true statement d genuinely undermining e. Since these philosophers want to illustrate an uncontroversial instance of knowledge as undefeated justified true belief, they present an example that is almost a conceptual necessity in their universe. Therefore, they confidently believe that e is an undefeated justification. However, consider that e is actually undermined by the true statement d* which can only be imagined or known through the mindset of the three-dimensional universe. In this case, it does not seem plausible to hold that d* is a genuine defeater by the standards of our hypothetical universe. We cannot reasonably expect these epistemologists to conceive of d* in the same way we do. Of course, one might say that d* is a genuine defeater both for the imagined philosophers' and S' (the agent in the philosophers' thought experiment) beliefs because what matters are the standards of our own universe. However, keep in mind that this example is just an imperfect analogy that is supposed to show how the lack of a sufficiently broad perspective in an epistemic paradigm might cause trouble within the very same paradigm. Similarly, if we suffer from the lack of a sufficiently broad perspective regarding unimaginable or unknowable statements, that situation would cause the same type of problem within our own universe. Therefore, I believe it makes sense to say that a true statement ought to meet the AC by *the standards of our universe* to be plausibly endorsed as a genuine defeater in our epistemic paradigm.

If the AC is not met, how a true statement can undermine a justification is not intelligible. At least some epistemic agents should be able to have a full conception of a true statement under optimal conditions so that it becomes clear why that true statement is incompatible with the justification. Otherwise, there would be no ground to hold that the statement undermines the justification. In other words, a broader perspective that makes an epistemically competent evaluation of available evidence does not exist when

we consider an unimaginable statement as a defeater candidate.

As I define the AC in relation to the knowability of a proposition, one might claim that defeasibility theories would suffer from circularity if they were combined with the AC (Klein, 1984, p. 228). This is because defeaters are employed in the definition of knowledge. If defeaters were also supposed to meet the AC, which is partially defined in terms of knowability, then the definition of knowledge would be linked with a standard that already contains a knowability requirement. Following from this, the defeasibility account of knowledge would be circular. However, this objection conflates the instance of knowledge within the experience of a particular epistemic agent (e.g., S in the thought experiment) with knowability of a proposition on the level of ideal epistemic agents (e.g., epistemologists who construct the thought experiment). Once one makes the distinction between these two levels, a sort of recursive definition would replace the apparent circularity. In other words, an actual instance of knowledge would be defined in terms of a potential instance of knowledge with different epistemic agents involved. In a thought experiment, S knows that p if her evidence e supports p and there is no genuine defeater d undermining e. d can be a genuine defeater only if the philosopher, who constructs the thought experiment, can potentially know that d is true (i.e., has an undefeated justified true belief that d is true). To put it another way, the philosopher cannot possibly disqualify S' belief on the grounds of a true statement which the former cannot possibly know.

5. Undefeated Justifications of Incompatible Metaphysical Beliefs

Both p and ¬ p cannot be true. Moreover, it is not possible for S to know that p and ¬ p because knowledge can only be attributed to true propositions. If one considers Klein's theory of defeasibility in connection with epistemically incompatible beliefs, it seems reasonable to hold that this incompatibility between beliefs has an impact on the undefeated justification condition as well. The strong defeasibility condition allegedly entails truth. An undefeated justification always leads to a true belief. If both p and ¬ p cannot be true, there must be a similar incompatibility between their justifications. Either the justification of p or the justification of ¬ p can be undefeated because it is not possible for both justifications to lead to true beliefs.

However, I believe that certain metaphysical beliefs and their negations can be justified in an undefeated way. The reason for this is that these justifications have no genuine defeaters that meet the AC. The controversy between this kind of competing justifications stems from our dissent regarding intuitions about specific notions. These intuitions do not seem to be empirically falsifiable or logically refutable. They are non-inferential beliefs, or dispositions to believe, which do seem true to the respective epistemic agents *solely because of how these agents understand the relevant notions* (Bonjour, 1998, p. 102; Russell, 2017). Since the justificatory role of competing intuitions is inherently tied to how they are non-inferentially understood by epistemic agents, we lack a sufficiently broad perspective to know if a statement, which might possibly undermine one of the competing justifications, is true. This is because, whenever we try to imagine a relevant true statement as a defeater, we either end up with our own

intuitions and discredit the opponents' intuitions or assume the opponents' intuitions to be true.[2] The first case is problematic since it circularly takes one party's controversial intuitions for granted. The second case does not provide a defeater either. This is because imagining the intellectual opponents' intuitions to be true would be in conflict with our first order epistemic attitudes. In such a case, there seems to be no reason to abandon our first order epistemic attitudes. As a result, in both cases, we are bound by the current set of intuitions and lack a broader epistemic perspective, which is conducive to successfully dissolving the conflict of intuitions. Although there is no conceivable ground to defeat one of the competing justifications in favor of the other, this does not mean that we know that both of these beliefs are true because both a belief and its negation cannot be true. This means that there is at least one possible undefeated justification that does not entail truth. Therefore, it follows that undefeated justified belief does not suffice to be counted as knowledge. The truth condition is needed to explain the difference between knowledge and undefeated justified metaphysical beliefs. I will explicate this argument step by step by illustrating it through a specific example. My example is the metaphysical debate on the first-cause argument about the emergence of the universe (cosmos).

5.1 A brief review of the first-cause debate

The first-cause argument implies the idea that the cosmos (universe) has an initial cause outside it (Taliaferro, 2014). When you consider the universe in its present state, it seems straightforward that the present state of the universe does not necessarily exist. Moreover, it *prima facie* makes sense to ask why the present state of the universe exists rather than another state since many of our scientific studies also rely on this question. To answer the question "what caused the present state of the universe?", we refer to the causal links between the present state and the previous state of the universe. Similarly, it makes sense to ask a similar question about the previous state of the universe. The answer would be that the previous state of the universe was caused by the state of the universe that existed before the previous state. The chain of past causes of the universe can go on infinitely in this way.

However, the proponents of the first-cause argument, who are usually theist philosophers of religion, contend that the idea of an infinite regress does not make sense since it leaves unanswered the question on the emergence of the universe (Taliaferro, 2014). According to these contenders, referring to the previous states of the universe is only a partial and very limited explanation. Although asking a question about the genesis of the universe as a whole is ostensibly intelligible, the appeal to the infinite regress of past causes does not provide a plausible answer to this question. Consider that you get a pencil from A, who got it from B, and so on indefinitely (Taliaferro, 2014). Referring to the past owners of the owners of the pencil successively does not provide an answer to the question "where does this pencil come from?" This explanation never

[2] Suspension of judgement might be considered to be the third option. However, it does not seem to influence the logic of my argument since agnosticism is considered to be just another competing position within the debate at hand.

arrives at an ultimate conclusion about the pencil's origin although it seems intelligible to ask that question. Since the idea of the first-cause allegedly generates a satisfactory explanation for an intelligible question (how does the universe exist?), many theist philosophers believe that appeal to the first-cause account is a more reasonable explanation than the infinite regress story.

Opponents of the first-cause argument, who are usually atheist philosophers, are quite skeptical about attributing a cause to the universe as a whole:

> What about the universe as a whole – could that have a cause? There is a straightforward reason for saying that the universe as a whole could not have a cause. Recall that the phrase 'the universe' is here being used to include space and time as well as matter. This means that there could not have been an event preceding the universe and bringing it about, for the simple reason that there was no time before the start of the universe in which that event could have occurred. (Everitt, 2004 , p. 70)

Everitt's claim simply relies on our intuitions about the notion of causation. Since causal relationships between things are embedded in our spatio-temporally defined universe, he holds that the very same notion of causation cannot be applied to external questions (i.e., questions that transcend the notion of space-time) about the universe as a whole. If the notion of causation is not applicable to our metaphysical questions regarding the universe, then asking what the ultimate cause of the universe is, is seriously confused. Once the plausibility of the question on the ultimate cause of the universe is undermined, there seems to be no reason to expect this question to have a satisfactory answer.

5.2 The source of the controversy

It is possible that both theistic and atheistic accounts of cosmology include mistakes in their theoretical justifications. For instance, a theist (an atheist) philosopher may have argued that there is (not) an ultimate cause of the universe by relying on bad reasons. In this sense, I do not say that any justification of these metaphysical beliefs are undefeated or qualified. However, it is also possible that some justifications of theistic and atheistic beliefs regarding the first-cause debate are undefeated in the sense that they include neither false empirical beliefs nor logical fallacies. This is possible when the contention between two beliefs merely relies on the conflicts of intuitions about certain key concepts. Let me propose the following thought experiment to illustrate my point.

Contemporary philosophers of religion often have a consensus on the reliability of advanced scientific findings and the rules of logic and argumentation. For instance, the Big Bang theory is recognized by atheist philosophers although, according to theists, this theory supports the first-cause argument. Similarly, both sides rely on the common rules of logic in their argumentation and refutation of their opponents' position. Imagine that the empirical findings and theoretical vision of the contemporary sciences are completely justified without any genuine defeater. Now, consider a group of philosophers, which includes both atheist and theist philosophers of religion. It is also possible that these philosophers do not commit any logical fallacies in their

argumentation regarding the first-cause argument. When you consider this hypothetical setting, the debate between atheist and theist philosophers on the first-cause argument is by no means unintelligible. Even if the justifications of both parties do not include any empirical mistakes or logical fallacies, their disagreement may still continue because of their conflicting intuitions about certain key concepts. The disagreement between two parties may merely rest upon a conflict about the notion of causation. Since atheists understand the notion of causation as a strictly naturalistic term, they disqualify the question on the ultimate cause of the universe.

Conversely, the proponents of the first-cause argument try to capture some other intuitions about the notion of causation. According to them, it is obvious that every contingent thing has a cause, which is prior to that thing. Following this, they see no reason to make an exception for the universe as a whole, which also contingently exists. Moreover, theists may hold that the spatio-temporal understanding of causation does not necessarily undermine that there may be quasi-causation on the level of the first-cause. Even if we do not exactly know what *kind of thing* this quasi-causation is, it can be coherently assumed that it is responsible for the emergence of the universe. Theists believe that the possibility of quasi-causation on the level of the meta-cosmos is more reasonable than assuming a contingent universe without a cause or a reason to exist. In this hypothetical setting, both parties try to put the burden of proof on the other side. Their disagreement relies on the intuitions that they have about causation and cosmos rather than on the empirical falsity or logical fallacy of one party.

One might claim that the disagreement about the notion of causation is not epistemic in the first place, since atheist and theist philosophers do talk past each other. The disagreement is semantic, as they have different concepts of causation. Thus, it has no relevance for epistemic properties. I believe that this objection overlooks the famous concept-conception distinction (Higginbotham, 1998; Ezcurdia, 1998). According to this distinction, the conditions for possessing a concept ought to be distinguished from the conditions for mastery of a concept (Ezcurdia, 1998, p. 188). Possessing a concept requires one to effectively employ the concept in one's communication with others. Unless the communication seriously breaks down in the intersubjective domain of epistemic agents, there is some room for disagreement that cannot be reduced to a mere equivocation. In contrast, concept mastery or having an adequate conception refers to applying the concept competently and/or using the concept within the true range of applicability. Following from this, a disagreement about how to conceive a concept is a disagreement of conception if there is considerable common ground in the first place. Given this common ground and the possibility of effective communication, epistemic agents can disagree about different properties or the range of reference in which the concept can be applied without talking past each other.

I believe that it makes sense to interpret the conflicts of intuition about the notion of causation as a disagreement about the true conception of causations rather than equivocation. First, atheist and theist philosophers' uses of the notion of causation do not have countless diverging properties. Instead, they essentially disagree on a single

important dimension of the concept: the question of whether causal relationships are necessarily embedded *within* the spatio-temporally defined universe. Hence, one can say that they still have considerable common ground about how to understand the concept of causation. This shared understanding makes it possible to hold that the disagreement is mainly related to the true range of applicability of a concept. This is a controversy of conception, not concept. Secondly, much of the philosophical literature is full of similar kinds of disagreements. Philosophical debate mainly starts with a minimally shared understanding of a concept and presents arguments to defend a particular conception that claims to correctly apply the concept. If one disqualifies the atheism-theism debate on the first-cause argument as equivocation, I believe this would lead to the absurd conclusion that most philosophical discussions are merely philosophers talking past each other.

5.3 Conflicting intuitions and the AC

Intuitions are categorized as a sort of belief or disposition to believe that X is true (Williamson, 2004, p. 126). My aim is not to argue for a specific definition of intuition. Rather, for the sake of the argument, I will assume that intuitions are non-inferential dispositions to believe *something* that are linked with a particular propositional content. Intuitions are highly important for our epistemic attitudes towards philosophical justifications. Philosophical accounts are usually challenged by the counter-examples, which indicate that these accounts are not in harmony with our strong intuitions about the respective subject matter. In the previous section, I illustrated an example in which it is possible for the proponents of two incompatible metaphysical beliefs to be in consensus about scientific facts and the rules of logic but not about certain intuitions regarding the key concepts of the disagreement at hand. Let me call the claim "the question on the ultimate cause of the universe is intelligible" p and its atheistic negation ¬ p. Suppose that q and ¬ q respectively represent conflicting intuitions such as the idea of causation (or quasi-causation) which is applicable to the universe as a whole and the atheistic negation of this proposition.

It is important to note that not every intuition constitutes good evidence to accept that a particular belief is true. There are many intuitive propositions that we give up after a period of careful deliberation or rational discussion. To avoid such cases, I restrain the scope of my discussion of intuitions by focusing on what Bonjour (1998) calls "rational insight". According to Bonjour, rational insights are intuitions that are persistent even after a period of consideration *"with a reasonable degree of care"* (1998, p. 114). This condition requires that the intuition is not a product of the lack of critical reflection. Moreover, rational insights are *a priori justifications* on the ground that their evidential power comes *solely from how their propositional contents are understood as necessary truths by epistemic agents* (Bonjour, 1998, pp. 102, 114). For intuitions to count as good evidence (rational insight), they ought to be accepted as necessarily true because of the way that

their propositional contents are inevitably understood.[3] It is not sufficient to say that intuitions seem self-evident.

I believe that both q and ¬ q meet the conditions specified by Bonjour. Firstly, we do not have any reason to assume that philosophers of religion are unreflective about the fundamental notions that they employ. Their philosophical work is assessed by the same standards we use for the work of other philosophers. If they were unreflective about the philosophical questions on which they work, it would be unlikely for them to be a part of the philosophical community. Moreover, atheist and theist philosophers are often in a scholarly dialogue with each other, forcing them to be even more reflective about their intuitions. Hence, the conflict of intuitions regarding the notion of causation does not seem to be the product of a lack of deliberation. Secondly, their intuitions about the notion of causation are firmly tied to how they necessarily understand the propositional content implied by their intuitions. For instance, an atheist philosopher of religion believes that her position is necessarily true because she understands the notion of causation as a naturalistic phenomenon that is not intelligible beyond the boundaries of the universe. Her intuitions are connected to a sense of metaphysical necessity about causation, a necessity derived from the very nature of causal relationships as it is understood by her.[4] This necessity makes them understand the notion of causation only in one particular way. Therefore, this intuition is different from other non-inferential beliefs (e.g., consider a non-inferential belief which is a product of wishful thinking), which seem to be true without deriving its persuasive power from how a certain concept is necessarily understood.

Let me now return to my example of incompatible metaphysical beliefs (p and ¬ p) and the corresponding intuitions (q and ¬ q). Since q and ¬ q express propositional content, it makes sense to assume that they are either true or false. However, logically, only one of them can be true. At this point, one can argue that both p and ¬ p cannot have undefeated justification because either q or ¬ q is false. However, I believe that q (or ¬ q) is not a genuine defeater of the justification of ¬ p (or p) because the AC is not met. The truth of q (or ¬ q) is by no means accessible to epistemic agents with normal rational capabilities under optimal conditions. Let us recall what I meant by accessibility in the AC: ideal epistemic agents ought to have a sufficiently broad perspective to be able to first imagine and then potentially know that d (a statement that defeats the initial justification) is true.

Now consider the following case: an atheist philosopher tries to come up with a thought

[3] Intuitions or rational insights are inevitable in belief-formation because even the most basic principles of reasoning or mathematical axioms rely on shared insights (Bonjour, 1998, p. 133). Similarly, intuitions seem to be inescapable in philosophical justifications (Williamson, 2007, p. 214). Nonetheless, there are some who hold that intuitions are unreliable for a variety of reasons such as the lack of consensus with regards to certain intuitions or the possibility that intuitions are the products of cognitive biases (Nagel, 2007, p. 799; Pust, 2016). However, a detailed discussion of intuitions is not within the scope of this paper. For a detailed discussion and the responses to the aforementioned objections, see (Bonjour, 1998, ch. 5).

[4] Defining metaphysical necessity is a controversial task. For the sake of my argument, I assume metaphysical necessity to be a type of necessity, which is derived from the nature or essence of a thing.

experiment that illustrates a defeater for either q or ¬ q. First, imagine that she assumes ¬ q to be true in her thought experiment. By assuming ¬ q to be true, she hypothetically realizes a defeater that refutes the belief that p. However, *assuming* ¬ q to be true and ¬ q's being *actually* true are two different things. Given the disagreement between atheist and theist philosophers, we do not have uncontroversial access to the truth of ¬ q. The atheist philosopher might hold that her rational insight that ¬ q guarantees that the statement is really true and hence it is a genuine defeater. However, the intuition ¬ q does not belong to a broader epistemic perspective in comparison to the intuition q. Conversely, q and ¬ q are epistemic opponents on the same level within an epistemic paradigm. Even if the atheist philosopher can create the fiction that ¬ q is true in her mind, this does not imply that she could potentially know that ¬ q. If she had some conclusive reasons that make her know that ¬ q is true in her thought experiment, then she would also be able to resolve the disagreement between atheists and theists in the real world. However, this is not the case. This fact demonstrates that atheist philosopher's epistemic perspective is not superior to or broader than that of her theist counterpart.

The second scenario is that the atheist philosopher assumes q to be true. In other words, she takes a statement, which she normally does not believe to be true, and assumes it to be true. Since imagination always brings a belief-like mental state about things that we normally do not believe, I assume that the atheist philosopher can create this type of mental state despite my skepticism. In this scenario, the atheist philosopher constructs the thought experiment to find a defeater against her own justification of ¬ p. There is, however, a clear problem with this scenario. Such an instance of epistemic imagination is not coherent because there is a conflict between the atheist philosophers' actual epistemic attitudes and what she assumes to be true. Furthermore, there seems to be no reason to attribute any evidential weight to her assumption. This is because she is still not able to show how her hypothetical epistemic outlook belongs to a broader perspective than her actual epistemic attitudes. Her assuming that q is true is based on neither intuition (rational insight) nor empirical evidence. Following from this, she is unable to construct a thought experiment implying a broader epistemic perspective that puts her in a position of evidential superiority. Hence, the AC is not met. Given our epistemic paradigm, the relevant intuitions are not possibly defeasible. We lack a sufficiently broad epistemic perspective to even imagine plausible thought experiments that could bring about plausible defeaters, let alone being able to find actual evidence to undermine the initial justification. Whenever we imagine a particular case, we suffer from a lack of sound reasons explaining why we assume a particular proposition (q or ¬ q) to be true.

To summarize, the AC refers to a perspective that makes possible epistemically competent evaluation of available evidence. Any attempt to qualify q as a genuine defeater should position itself on a superior epistemic level to be able to evaluate available evidence. However, assuming that q is true is not a matter of broader perspective formation from which we have more information and are able to construct a Tom-Grabit-like thought experiment. This is rather a counter-position, which is on the

same level with the proponents of ¬ q. There is no reasonable ground to assume that q is true except for the fact that q non-inferentially seems to be true to some people as a rational insight. Similarly, ¬ q non-inferentially seems to be true to some other group of people. Both q and ¬ q are not empirically falsifiable or logically refutable. Hence, I believe that no epistemic agent is capable of constructing a Tom-Grabit-like thought experiment since the truth of q (or ¬ q) is not fully conceivable from a broader perspective that makes an epistemically competent evaluation of evidence possible. A competent epistemic evaluation of these intuitions has no grounds. Since the truth of q (or ¬ q) is not accessible to epistemic agents under any (optimal) condition, it is not plausible to believe that these justifications can have genuine defeaters, which are intelligible through our thought experiments.

One might claim that we should not accept the justificatory role of q and ¬ q in the first place if we concede that a competent epistemic evaluation of them is not possible. In other words, we cannot legitimately hold these intuitions, if we do not have access to their truths. By delineating what I mean by "competent epistemic evaluation", I hope to provide a satisfactory response to this criticism. To be explicit, I mean a process of evaluation that conclusively ends the divergence of opinion among different epistemic agents. For instance, when we assume ourselves to be operating under a broader epistemic perspective in the Tom Grabit example, our knowledge that Tom Grabit has a twin brother is conclusive. If S (the epistemic agent in the thought experiment) had the opportunity to talk to us (the agents holding the broader epistemic perspective), she would immediately change her mind once we reveal this piece of information to her (namely, that Tom Grabit has a twin). Hence, a broader epistemic perspective makes it possible to conclusively change the other agent's mind. It brings about a certainty-like situation that dissolves the *prima facie* disagreement. However, for the aforementioned reasons, q and ¬ q cannot be subject to a similar kind of epistemic evaluation that might conclusively end the disagreement among epistemic agents. Nonetheless, q and ¬ q can still constitute a kind of evidence that support metaphysical beliefs especially for the proponents of fallibilism, who believe that beliefs cannot be conclusively justified. q and ¬ q are rational insights whose evidential weight comes from the way they are necessarily understood by epistemic agents. As they are rational insights, they do not need to rely on further evidence before they play a role in a justification. Requiring that we should be certain regarding their truth (hence should have no insolvable controversy) before we accept their justificatory power is problematic, since this assumption expels the whole class of disagreement-driven philosophical argumentation from the domain of justification. Not only are metaphysical beliefs about the origin of the universe based on a divergence of intuitions, but also, so are any kind of philosophical subject containing competing conceptions of certain key notions. Even if achieving certainty is necessary to conclusively defeat this type of justifications, it would be too demanding to require that a reflective endorsement of philosophical beliefs is supposed to entail certainty in order to be accepted as a genuine justification. If rational epistemic agents do persistently hold these beliefs after a reasonable period of reflection and dialogue with intellectual opponents, it seems sensible to accept these beliefs as plausible epistemic reasons despite the lack of certainty.

Moreover, the justificatory role of q and ¬ q is even compatible with some forms of infallibilism (e.g., defeasibility theories). For instance, defeasibility theorists might still hold that the evaluation of justifications is strictly subject to their infallibilist standards when it comes to the components of a justification related to empirical evidence and the rules of logic. Whereas controversial philosophical and/or metaphysical beliefs based on rational insights can potentially have undefeated but fallible justifications. In the model that I offer, defeasibility theories function as hybrid systems that implement both fallible and infallible understandings of justification in different domains of belief-formation. The main contribution of my account is to show how our conventional understanding of indefeasibility can be made compatible with the disagreement-driven (and fallible) nature of philosophical and/or metaphysical justification.

5.4 Undefeated justification without knowledge

Since proponents and opponents of the first-cause argument agree on the findings of contemporary science, it can be said that their sets of evidence have some common components. For the example I have been discussing, assume e is one said common component. The conjunction (e & q) justifies p whereas the conjunction (e & ¬ q) justifies ¬ p. Both justifications have no genuine defeaters because no true statement, which can be assumed to be known, undermines these justifications. In this sense, (e & q) and (e & ¬ q) are undefeated justifications.

Given that both p and ¬ p cannot be true, there is at least one undefeated justification that does not entail truth. If p is true and (e & q) entails truth, then it is not possible for (e & ¬ q) to entail truth because ¬ p would be false. Under these circumstances, there is also at least one undefeated justified belief that cannot be an instance of knowledge. Explicitly, this is the case because knowledge entails truth.

5.5 The truth condition is needed in the analysis of knowledge

Since undefeated justification does not always guarantee truth, there must be a gap between knowledge and undefeated justified belief. Let me explicate this below in light of my example on metaphysical beliefs:

1. (e & q) is the undefeated justification of p.
2. (e & ¬ q) is the undefeated justification of ¬ p.
3. Both p and ¬ p cannot be true.
4. Either (e & q) or (e & ¬ q) do not entail truth.
5. There is at least one undefeated justification that does not entail truth.
6. Knowledge entails truth.
7. Undefeated justified belief does not suffice to define knowledge.

The case of metaphysical beliefs shows that it is possible for some beliefs to have undefeated justifications, which do not entail truth. The incapacity of these undefeated justifications to entail truth stems from the fact that a polar opposite belief can also have an undefeated justification. However, both a proposition and its negation cannot be true. Given that knowledge entails truth, the formulation of undefeated justified belief does not suffice to be the proper definition of knowledge. Following this, I conclude that a distinct truth condition is needed in the analysis of knowledge.

6. Conclusion

I argued that the truth condition is not superfluous in the analysis of knowledge because undefeated justification does not always entail truth. There are some possible justifications for incompatible metaphysical beliefs (p and ¬ p), which have no genuine defeaters. The incongruity between the justifications rests upon a conflict of intuitions. However, the truths of these intuitions are not accessible to epistemic agents under optimal conditions. A broader perspective from which the falsity of an intuition can be shown is inconceivable. Therefore, there is no intelligible way of imagining or knowing genuine defeaters for these justifications. Given that both p and ¬ p cannot be true and knowledge entails truth, I claimed that there is at least one undefeated justification that does not entail truth. Given this independence between justification and truth, a distinct truth condition is essential in the analysis of knowledge.

Acknowledgement

I would like to thank Gordian Haas, Enzo Rossi, Cenk Özdağ, Kaan Tabakçı, and Danielle Scheil for many helpful suggestions. I would also like to thank the anonymous reviewers whose criticisms helped me improve the initial draft of this paper.

REFERENCES

Bonjour, L. (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.

Everitt, N. (2004). *The Non-Existence of God*. London: Routledge.

Ezcurdia, M. (1998). The Concept-Conception Distinction. *Philosophical Issues*, *9*, 187-192.

Gettier, E. L. (1963). Is Justified True Belief Knowledge*? Analysis, 23*(6), 121-123.

Hetherington, S. (n.d.). Gettier Problems. In *Internet Encyclopedia of Philosophy* (J. Fieser & B. Dowden, Eds.). Retrieved from https://www.iep.utm.edu/gettier

Higginbotham, J. (1998). Conceptual Competence. *Philosophical Issues*, *9*, 149-162.

Klein, P. D. (1980). Misleading Evidence and the Restoration of Justification. *Philosophical Studies*, *37*(1), 81-89.

Klein, P. D. (1984). *Certainty: A Refutation of Scepticism*. Minneapolis: University of Minnesota Press.

Lehrer, K. & Paxson, T. (1969). Knowledge: Undefeated Justified True Belief. *Journal of Philosophy 66*(8), 225-237.

Lehrer, K. (1990). *Theory of Knowledge*. Boulder and Colorado: Westview Press.

Levy, S. R. (1977). Defeasibility Theories of Knowledge. *Canadian Journal of Philosophy, 7*(1), 115-123.

Merricks, T. (1995). Warrant Entails Truth. *Philosophy and Phenomenological Research, 55*, 841-855.

Nagel, J. (2007). Epistemic Intuitions. *Philosophy Compass, 2*(6), 792-819.

Pust, J. (2016). Intuition. In *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved from https://plato.stanford.edu/archives/spr2016/entries/intuition

Russell, B. (2017). *A Priori* Justification and Knowledge. In *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved from https://plato.stanford.edu/archives/sum2017/entries/apriori

Sudduth, M. (n.d.). Defeaters in Epistemology. In *Internet Encyclopedia of Philosophy* (J. Fieser & B. Dowden, Eds.). Retrieved from https://www.iep.utm.edu/ep-defea

Taliaferro, C. (2014). Philosophy of Religion. In *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved from http://plato.stanford.edu/archives/win2014/entries/philosophy-religion

Williamson, T. (2004). Philosophical 'Intuitions' and Scepticism about Judgement. *Dialectica, 58*(1), 109-153.

Williamson, T. (2007). *Philosophy of Philosophy*. Oxford: Blackwell Publishing.

Zagzebski, L. (1994). The Inescapability of Gettier Problems. *Philosophical Quarterly, 44*(174), 65-73.